PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/40, 15/11, C07K 14/08, C12P 21/02, 19/34, C12Q 1/70, G01N 33/569, A61K 39/12, 39/42, 48/00

A1

(11) International Publication Number: **WO 96/40932**

(43) International Publication Date: 19 December 1996 (19.12.96)

(21) International Application Number: PCT/US96/08962

(22) International Filing Date: 7 June 1996 (07.06.96)

(30) Priority Data:
08/478,316 7 June 1995 (07.06.95) US

(71)(72) Applicants and Inventors: PAUL, Prem, S. [US/US]; 4206 Arizona Circle, Ames, IA 50014 (US). MENG, Xiang-Jin [CN/US]; 725 Pammel Court, Ames, IA 50014 (US). HALBUR, Patrick [US/US]; 3211 Kingman Road, Ames, IA 50014 (US). MOROZOV, Igor [RU/US]; Apartment #156, 407 South 5th Street, Ames, IA 50010 (US).

(74) Agents: LAVALLEYE, Jean-Paul et al.; Oblon, Spivak, McClelland, Maier & Neustadt, P.C., Crystal Square Five, 4th floor, 1755 Jefferson Davis Highway, Arlington, VA 22202 (US).

(81) Designated States: CA, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: VACCINE FOR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS

(57) Abstract

The present invention provides a purified preparation containing, for example, a polynucleic acid encoding at least one polypeptide selected from the group consisting of proteins encoded by one or more open reading frames (ORF's) of an Iowa strain of porcine reproductive and respiratory syndrome virus (PRRSV), antigenic regions of such proteins which are at least 5 amino acids in length and which effectively protect a porcine host against a subsequent challenge with a PRRSV isolate, and combinations thereof in which amino acids non-essential for antigenicity may be conservatively substituted. The present invention also concerns a polypeptide encoded by such a polynucleic acid; a vaccine comprising an effective amount of such a polynucleic acid or protein; antibodies which specifically bind to such a polynucleic acid or protein; methods of producing the same; and methods of protecting a pig against a PRRSV, treating a pig infected by a PRRSV, and detecting a PRRSV using the same.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

VACCINE FOR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS

This is a continuation-in-part of application Serial No. 08/301,435, filed on September 1, 1994, pending, which is a continuation-in-part of application Serial No. 08/131,625, filed on October 5, 1993, pending, which is a continuation-in-part of application Serial No. 07/969,071, filed on October 30, 1992, now abandoned. The entire contents of application Serial Nos. 08/301,435 and 08/131,625, filed on September 1, 1994 and October 5, 1993, respectively, are incorporated herein by reference in their entireties.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention concerns polynucleic acids isolated from a porcine reproductive and respiratory syndrome virus (PRRSV), a protein and/or a polypeptide encoded by the polynucleic acids, a vaccine which protects pigs from a PRRSV based on the protein or polynucleic acids, methods of making the proteins, polypeptides and polynucleic acids, a method of protecting a pig from PRRS using the vaccine, a method of producing the vaccine, a method of treating a pig infected by or exposed to a PRRSV, and a method of detecting a PRRSV.

Discussion of the Background:

Porcine reproductive and respiratory syndrome (PRRS), a new and severe disease in swine, was first reported in the U.S.A. in 1987, and was rapidly recognized in many western European countries (reviewed by Goyal, J. Vet. Diagn. Invest., 1993, 5:656-664; and in U.S. Application Serial Nos. 08/131,625 and 08/301,435). The disease is characterized by reproductive failure in sows and gilts, pneumonia in young growing pigs, and an increase in preweaning mortality (Wensvoort et al., Vet. Q., 13:121-130, 1991; Christianson et al., 1992, Am. J. Vet. Res. 53:485-488; U.S. Application Serial Nos. 08/131,625 and 08/301,435).

The causative agent of PRRS, porcine reproductive and respiratory syndrome virus (PRRSV), was identified first in Europe and then in the U.S.A. (Collins et al., 1992, J. Vet. Diagn. Invest., 4:117-126). The European strain of PRRSV, designated as Lelystad virus (LV), has been cloned and sequenced (Meulenberg et al., 1993, Virology, 192:62-72 and J. Gen. Virol., 74:1697-1701; Conzelmann et al., 1993, Virology, 193:329-339).

PRRSV was provisionally classified in the proposed new virus family of *Arteriviridae*, which includes equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV) and simian hemorrhagic fever virus (SHFV) (Plagemann and Moennig, 1992, Adv. Virus. Res., 41:99-192; Godeny et al., 1993, Virology, 194:585-596; U.S. Application Serial Nos. 08/131,625

and 08/301,435). This group of single plus-strand RNA viruses shares many characteristics such as genome organization, replication strategy, morphology and macrophage tropism (Meulenberg et al., 1993; U.S. Application Serial Nos. 08/131,625 and 08/301,435). Subclinical infections and persistent viremia with concurrent antibody production are also characteristic histopathologic properties of the arteriviruses.

Antigenic, genetic and pathogenic variations have been reported among PRRSV isolates (Wensvoort et al., 1992, J. Vet. Diagn. Invest., 4:134-138; Mardassi et al., 1994, J. Gen. Virol., 75:681-685; U.S. Application Serial Nos. 08/131,625 and 08/301,435). Furthermore, U.S. and European PRRSV represent two distinct genotypes (U.S. Application Serial Nos. 08/131,625 and 08/301,435). Antigenic variability also exists among different North American isolates as well (Wensvoort et al., 1992). Marked differences in pathogenicity have been demonstrated not only between U.S. and European isolates, but also among different U.S. isolates (U.S. Application Serial Nos. 08/131,625 and 08/301,435).

The genomic organization of arteriviruses resembles coronaviruses and toroviruses in that their replication involves the formation of a 3'-coterminal nested set of subgenomic mRNAs (sg mRNAs) (Chen et al., 1993, J. Gen. Virol. 74:643-660; Den Boon et al., 1990, J. Virol., 65:2910-2920; De Vries et al., 1990, Nucleic Acids Res., 18:3241-3247; Kuo et

al., 1991, J. Virol., 65:5118-5123; Kuo et al., 1992; U.S. Application Serial Nos. 08/131,625 and 08/301,435). Partial sequences of several North American isolates have also been determined (U.S. Application Serial Nos. 08/131,625 and 08/301,435; Mardassi et al., 1994, J. Gen. Virol., 75:681-685).

The genome of PRRSV is polyadenylated, about 15 kb in length and contains eight open reading frames (ORFs; Meulenberg et al., 1993; U.S. Application Serial Nos. 08/131,625 and 08/301,435). ORFs 1a and 1b probably encode viral RNA polymerase (Meulenberg et al., 1993). ORFs 5, 6 and 7 were found to encode a glycosylated membrane protein (E), an unglycosylated membrane protein (M) and a nucleocapsid protein (N), respectively (Meulenberg et al., 1995). ORFs 2 to 4 appear to have the characteristics of membrane-associated proteins (Meulenberg et al., 1993; U.S. Application Serial No. 08/301,435). However, the translation products of ORFs 2 to 4 were not detected in virus-infected cell lysates or virions (Meulenberg et al., 1995).

The major envelope glycoprotein of EAV encoded by ORF 5 may be the virus attachment protein, and neutralizing monoclonal antibodies (MAbs) are directed to this protein (de Vries, J. Virol. 1992; 66:6294-6303; Faaberg, J. Virol. 1995; 69:613-617). The primary envelope glycoprotein of LDV, a closely related member of PRRSV, is also encoded by ORF 5, and several different neutralizing MAbs were found to specifically

immunoprecipitate the ORF 5 protein (Cafruny et al., Vir. Res., 1986; 5:357-375). Therefore, it is likely that the major envelope protein of PRRSV encoded by ORF 5 may induce neutralizing antibodies against PRRSV.

It has been proposed that antigenic variation of viruses is the result of direct selection of variants by the host immune responses (reviewed by Domingo et al., J. Gen. Virol. 1993, 74:2039-2045). Thus, these hypervariable regions are likely due to the host immune selection pressure and may explain the observed antigenic diversity among PRRSV isolates.

The M and N proteins of U.S. PRRSV isolates, including ISU 3927, are highly conserved (U.S. Application Serial No. 08/301,435). The M and N proteins are integral to preserving the structure of PRRSV virions, and the N protein may be under strict functional constraints. Therefore, it is unlikely either that (a) the M and N proteins are subjected to major antibody selection pressure or that (b) ORFs 6 and 7, which are likely to encode the M and N proteins, are responsible for or correlated to viral virulence. Interestingly, however, higher sequence variation of the LDV M protein was observed between LDV isolates with differing neurovirulence (Kuo et al., 1992, *Vir. Res.* 23:55-72).

ORFs 1a and 1b are predicted to translate into a single protein (viral polymerase) by frameshifting. ORFs 2 to 6 may encode the viral membrane associated proteins.

In addition to the genomic RNA, many animal viruses produce one or more sg mRNA species to allow expression of viral genes in a regulated fashion. In cells infected with PRRSV, seven species of virus-specific mRNAs representing a 3'-coterminal nested set are synthesized (mRNAs 1 to 7, in decreasing order of size). mRNA 1 represents the genomic mRNA. Each of the sg mRNAs contains a leader sequence derived from the 5'-end of the viral genome.

The numbers of the sg mRNAs differ among arteriviruses and even among different isolates of the same virus. A nested set of 6 sg mRNAs was detected in EAV-infected cells and European PRRSV-infected cells. However, a nested set of six (LDV-C) or seven (LDV-P) sg mRNAs, in addition to the genomic RNA, is present in LDV-infected cells. The additional sg mRNA 1-1 of LDV-P contains the 3'-end of ORF 1b and can potentially be translated to a protein which represents the C-terminal end of the viral polymerase. Sequence analysis of the sg mRNAs of LDV and EAV indicates that the leader-mRNA junction motif is conserved. Recently, the leader-mRNA junction sequences of the European LV were also shown to contain a common motif, UCAACC, or a highly similar sequence.

The sg mRNAs have been shown to be packaged into the virions in some coronaviruses, such as bovine coronavirus (BCV) and transmissible gastroenteritis virus (TGEV). However, only trace amounts of the sg mRNAs were detected in purified virions of mouse hepatitis virus (MHV), another

coronavirus. The sg mRNAs of LDV, a closely related member of PRRSV, are also not packaged in the virions, and only the genomic RNA was detected in purified LDV virions.

The sg mRNAs of LDV and EAV have been characterized in detail. However, information regarding the sg mRNAs of PRRSV strains, especially the U.S. PRRSV, is very limited. Thus, a need is felt for a more thorough molecular characterization of the sg mRNAs of U.S. PRRSV.

The packaging signal of MHV is located in the 3'-end of ORF 1b, thus only the genomic RNA of MHV is packaged. The sg mRNAs of BCV and TGEV, however, are found in purified virions. The packaging signal of BCV and TGEV has not been determined. The Aura alphavirus sg mRNA is efficiently packaged into the virions, presumably because the packaging signal is present in the sg mRNA. The sindbis virus 26S sg mRNA is not packaged into virions because the packaging signal is located in the genome segment (not present in sg mRNA). The sg mRNAs of LDV, a closely related member of PRRSV, are also not packaged into the virions.

Many mechanisms are involved in the generation of the sg mRNAs. It has been proposed that coronaviruses utilize a unique leader RNA-primed transcription mechanism in which a leader RNA is transcribed from the 3' end of the genome-sized negative-stranded template RNA, dissociates from the template, and then rejoins the template RNA at downstream intergenic regions to prime the transcription of sg mRNAs. The model

predicts that the 5'-leader contains a specific sequence at its 3'-end which is repeated further downstream in the genome, preceding each of the ORFs 2 to 7. The leader joins to the body of each of the sg mRNAs via the leader-mRNA junction segment.

PRRSV is an important cause of pneumonia in nursery and weaned pigs. PRRSV causes significant economic losses from pneumonia in nursery pigs (the exact extent of which are not fully known). Reproductive disease was the predominant clinical outcome of PRRSV infections during the past few years, due to the early prevalence of relatively low virulence strains of PRRSV. Respiratory disease has now become the main problem associated with PRRSV, due to the increasing prevalence of relatively high virulence strains of PRRSV. A need is felt for a vaccine to protect against disease caused by the various strains of PRRSV.

Surprisingly, the market for animal vaccines in the U.S. and worldwide is larger than the market for human vaccines. Thus, there exists an economic incentive to develop new veterinary vaccines, in addition to the substantial public health benefit which is derived from protecting farm animals from disease.

SUMMARY OF THE INVENTION

Accordingly, one object of the present invention is to provide a polynucleic acid isolated from a porcine reproductive and respiratory syndrome virus (PRRSV).

It is a further object of the present invention to provide an isolated polynucleic acid which encodes a PRRSV protein.

It is a further object of the present invention to provide a PRRSV protein, either isolated from a PRRSV or encoded by a PRRSV polynucleic acid.

It is a further object of the present invention to provide a protein- or polynucleic acid-based vaccine which protects a pig against PRRS.

It is a further object of the present invention to provide a method of raising an effective immunological response against a PRRSV using the vaccine.

It is a further object of the present invention to provide a method of producing a protein- or polynucleic acid-based vaccine which protects a pig against PRRS.

It is a further object of the present invention to provide a method of treating a pig exposed to a PRRSV or suffering from PRRS.

It is a further object of the present invention to provide a method of detecting PRRSV.

It is a further object of the present invention to provide an antibody which immunologically binds to a PRRSV protein or to an antigenic region of such a protein.

It is a further object of the present invention to provide an antibody which immunologically binds to a protein- or polynucleic acid-based vaccine which protects a pig against a PRRSV.

It is a further object of the present invention to provide a diagnostic kit for assaying or detecting a PRRSV.

It is a further object of the present invention to provide the above objects, where the PRRS virus is an Iowa strain of PRRSV.

These and other objects, which will become apparent during the following description of the preferred embodiments, have been provided by a purified and/or isolated polypeptide selected from the group consisting of proteins encoded by one or more open reading frames (ORF's) of an Iowa strain of porcine reproductive and respiratory syndrome virus (PRRSV), proteins at least 94% but less than 100% homologous with a protein encoded by an ORF 2 of an Iowa strain of PRRSV, proteins at least 88% but less than 100% homologous with a protein encoded by ORF 3 of an Iowa strain of PRRSV, proteins at least 93% homologous with an ORF 4 of an Iowa strain of PRRSV, proteins at least 90% homologous with an ORF 5 of an Iowa strain of PRRSV, proteins at least 97% but less than 100% homologous with proteins encoded by one or both of ORF 6 and

ORF 7 of an Iowa strain of PRRSV, antigenic regions of such proteins which are at least 5 amino acids in length and which effectively stimulate protection in a porcine host against a subsequent challenge with a PRRSV isolate, and combinations thereof; an isolated polynucleic acid which encodes such a polypeptide or polypeptides; a vaccine comprising an effective amount of such a polynucleotide or polypeptide(s); antibodies which specifically bind to such a polynucleotide or polypeptide; methods of producing the same; and methods of (i) effectively protecting a pig against PRRS, (ii) treating a pig exposed to a PRRSV or suffering from PRRS, and (iii) detecting a PRRSV using the same.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows a nucleotide sequence comparison of ORFs 2 to 5 of U.S. isolates ISU 79, ISU 1894, ISU 3927, ISU 22 and ISU 55 with other known PRRSV isolates;

Figures 2A, 2B, 2C and 2D respectively show the alignment of the deduced amino acid sequences of ORF 2, ORF 3, ORF 4 and ORF 5 of U.S. isolates ISU 79, ISU 1894, ISU 22, ISU 55 and ISU 3927 with other known PRRSV isolates;

Fig. 3 shows a phylogenetic tree based on the nucleotide sequences of ORFs 2 to 7 of seven U.S. PRRSV isolates with differing virulence;

Figure 4 shows a Northern blot analysis of RNAs isolated from ISU 3927-infected CRL 11171 cells (lane 1) and from purified virions of ISU 3927 (lane 2);

Figure 5 shows a Northern blot analysis of total intracellular RNAs isolated from CRL 11171 cells infected with ISU22 (lane 1), ISU 55 (lane 2), ISU 79 (lane 3), ISU 1894 (lane 4) and ISU 3927 (lane 5), respectively;

Figures 6A and 6B show a Northern hybridization of total RNAs isolated from CRL 11171 cells infected with ISU 79 at different multiplicities of infection (m.o.i.) (A), and polyadenylated RNA from cells infected with PRRSV isolates ISU 55 and ISU 79 (B);

Figures 7A and 7B show a Northern blot analysis of total intracellular mRNAs isolated from CRL 11171 cells infected with ISU 1894 (A) and ISU 79 (B);

Figures 8A and 8B show RT-PCR amplification of the 5'-terminal sequences of the sg mRNAs 3 and 4 of ISU 1894 (lane 1) and sg mRNAs 3, 4 and 4-1 of ISU 79 (lane 2) (A) where lane L is a 1-kb marker; and the leader-mRNA junction sequences of sg mRNAs 3 and 4 of ISU 79 and ISU 1894 and of sg mRNA 4-1 of ISU 79 (B), where the locations of the leader-mRNA junction sequences in the genomes relative to the start codon of each ORF were indicated by minus (-) numbers of nucleotides upstream of the ORFs; and

Figure 9 shows the sequence alignment of ORFs 2 to 7 of ISU 1894 and ISU 79, where the start codon of each ORF is

indicated by +>, the termination codon of each ORF is indicated by asterisks (*), the determined or predicted leader-mRNA junction sequences are underlined and the locations of the leader-mRNA junction sequences relative to the start codon of each ORF are indicated by minus (-) numbers of nucleotides upstream of each ORF.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the present application, the nucleotide sequences of the ORFs 2 to 5 of a low virulence isolate and four other Iowa strain PRRSV isolates with "moderate" and high virulence have been determined. Based on comparisons of ORFs 2 to 7 of various PRRSV isolates, the least virulent U.S. isolate known (ISU 3927) has relatively high sequence variations in ORFs 2 to 4, as compared to the variations in other U.S. isolates. Furthermore, based on analysis of the sequences of the ORFs, at least three minor genotypes exist within the major genotype of U.S. PRRSV.

Sequence analysis of the ORF 5 protein of different PRRSV isolates reveal three hypervariable regions which contained non-conserved amino acid substitutions. These regions are hydrophilic and also antigenic as predicted by computer analysis.

In the present invention, a "porcine reproductive and respiratory syndrome virus" or "PRRSV" refers to a virus which causes the diseases PRRS, PEARS, SIRS, MSD and/or PIP (the

term "PIP" now appears to be disfavored), including the Iowa strain of PRRSV, other strains of PRRSV found in the United States (e.g., VR 2332), strains of PRRSV found in Canada (e.g., IAF-exp91), strains of PRRSV found in Europe (e.g., Lelystad virus, PRRSV-10), and closely-related variants of these viruses which may have appeared and which will appear in the future.

The "Iowa strain" of PRRSV includes (a) PRRSV isolates deposited in the American Type Culture Collection by the present inventors and/or described in this application and/or in either of prior U.S. Application Serial Nos. 08/131,625 and 08/301,435, (b) PRRS viruses which produce more than six sg mRNAs when cultured or passaged in CRL 11171 cells, (c) PRRSVs which produce at least 40% gross lung lesions or lung consolidation in 5-week-old caesarean-derived, colostrum-deprived piglets 10 days post-infection, (d) a PRRSV isolate having a genome which encodes a protein having the minimum homology to a PRRSV ORF described in Table 2 below, and/or (d) any PRRSV isolate having the identifying characteristics of such a virus.

The present vaccine is effective if it protects a pig against infection by a porcine reproductive and respiratory syndrome virus (PRRSV). A vaccine protects a pig against infection by a PRRSV if, after administration of the vaccine to one or more unaffected pigs, a subsequent challenge with a biologically pure virus isolate (e.g., VR 2385, VR 2386, or

other virus isolate described below) results in a lessened severity of any gross or histopathological changes (e.g., lesions in the lung) and/or of symptoms of the disease, as compared to those changes or symptoms typically caused by the isolate in similar pigs which are unprotected (i.e., relative to an appropriate control). More particularly, the present vaccine may be shown to be effective by administering the vaccine to one or more suitable pigs in need thereof, then after an appropriate length of time (e.g., 1-4 weeks), challenging with a large sample (10^{3-7} $TCID_{50}$) of a biologically pure PRRSV isolate. A blood sample is then drawn from the challenged pig after about one week, and an attempt to isolate the virus from the blood sample is then performed (e.g., see the virus isolation procedure exemplified in Experiment VIII below). Isolation of the virus is an indication that the vaccine may not be effective, and failure to isolate the virus is an indication that the vaccine may be effective.

Thus, the effectiveness of the present vaccine may be evaluated quantitatively (i.e., a decrease in the percentage of consolidated lung tissue as compared to an appropriate control group) or qualitatively (e.g., isolation of PRRSV from blood, detection of PRRSV antigen in a lung, tonsil or lymph node tissue sample by an immunoperoxidase assay method [described below], etc.). The symptoms of the porcine reproductive and respiratory disease may be evaluated quantitatively (e.g., temperature/ fever), semi-quantitatively

(e.g., severity of respiratory distress [explained in detail below], or qualitatively (e.g., the presence or absence of one or more symptoms or a reduction in severity of one or more symptoms, such as cyanosis, pneumonia, heart and/or brain lesions, etc.).

An unaffected pig is a pig which has either not been exposed to a porcine reproductive and respiratory disease infectious agent, or which has been exposed to a porcine reproductive and respiratory disease infectious agent but is not showing symptoms of the disease. An affected pig is one which shows symptoms of PRRS or from which PRRSV can be isolated.

The clinical signs or symptoms of PRRS may include lethargy, respiratory distress, "thumping" (forced expiration), fevers, roughened haircoats, sneezing, coughing, eye edema and occasionally conjunctivitis. Lesions may include gross and/or microscopic lung lesions, myocarditis, lymphadenitis, encephalitis and rhinitis. In addition, less virulent and non-virulent forms of PRRSV and of the Iowa strain have been found, which may cause either a subset of the above symptoms or no symptoms at all. Less virulent and non-virulent forms of PRRSV can be used according to the present invention to provide protection against porcine reproductive and respiratory diseases nonetheless.

The phrase "polynucleic acid" refers to RNA or DNA, as well as mRNA and cDNA corresponding to or complementary to the

RNA or DNA isolated from the virus or infectious agent. An "ORF" refers to an open reading frame, or polypeptide-encoding segment, isolated from a viral genome, including a PRRSV genome. In the present polynucleic acid, an ORF can be included in part (as a fragment) or in whole, and can overlap with the 5'- or 3'-sequence of an adjacent ORF (see for example, Fig. 1 and Experiment 1 below). A "polynucleotide" is equivalent to a polynucleic acid, but may define a distinct molecule or group of molecules (e.g., as a subset of a group of polynucleic acids).

In the Experiments described hereinbelow, the isolation, cloning and sequencing of ORFs 2 to 5 of (a) a low virulence U.S. PRRSV isolate and (b) two other U.S. PRRSV isolates of varying virulence were determined. The nucleotide and deduced amino acid sequences of these three U.S. isolates were compared with the corresponding sequences of other known PRRSV isolates (see, for example, U.S. Application Serial No. 08/301,435). The results indicate that considerable genetic variations exist not only between U.S. PRRSV and European PRRSV, but also among the U.S. isolates as well.

The amino acid sequence identity between the seven U.S. PRRSV isolates studied was 91-99% in ORF 2, 86-98% in ORF 3, 92-99% in ORF 4 and 88-97% in ORF 5. The least virulent U.S. isolate known (ISU 3927) has higher sequence variations in ORFs 2 to 4 than in ORFs 5 to 7, as compared to other U.S. isolates. Three hypervariable regions with antigenic

potential have been identified in the major envelope glycoprotein encoded by ORF 5.

Pairwise comparison of the sequences of ORFs 2 to 7 and phylogenetic tree analysis implied the existence of at least three groups of PRRSV variants (or minor genotypes) within the major genotype of U.S. PRRSV. The least virulent U.S. isolate known forms a distinct branch from other U.S. isolates with differing virulence. The results of this study have implications for the taxonomy of PRRSV and vaccine development.

In a further experiment, the sg mRNAs in PRRSV-infected cells were characterized. The data showed that a 3'-coterminal nested set of six or seven sg mRNAs is formed in cells infected with different isolates of PRRSV. However, unlike some of the coronaviruses and alphavirus, the sg mRNAs of PRRSV are not packaged into the virion, and only was the genomic RNA of PRRSV detected in purified virions. Variations in the numbers of the sg mRNAs among different PRRSV isolates with differing virulence were also observed. Further sequence analysis of ORFs 2 to 7 of two U.S. isolates and their comparison with the European LV reveal the heterogeneic nature of the leader-mRNA junction sequences of PRRSV.

As demonstrated in Experiment 2 below, a 3'-coterminal nested set of six or more sg mRNAs is formed in cells infected with different isolates of PRRSV. The presence of a nested set of sg mRNAs further indicates that U.S. PRRSV, like the

European isolate Lelystad virus (LV), belongs to the newly proposed *Arteriviridae* family including LDV, EAV and SHFV. Northern blot analysis with ORF-specific probes indicates that the structure of the PRRSV sg mRNAs is polycistronic, and each of the sg mRNAs except for sg mRNA 7 contains multiple ORFs. Therefore, the sequence of each sg mRNA is contained within the 3'-portion of the next larger sg mRNA, and not all 5'-ends of the sg mRNAs overlap with the sequences of the smaller sg mRNAs.

There is no apparent correlation, however, between the numbers of sg mRNAs and viral pneumovirulence. An additional species, sg mRNA 4-1, was found to contain a small ORF (ORF 4-1) with a coding capacity of 45 amino acids at its 5'-end.

In Experiment 2 below, the sg mRNAs of PRRSV are shown not to be packaged into the virions. Whether sg mRNAs are packaged into virions may depend an whether the sg mRNAs contain a packaging signal. Since the sg mRNAs of PRRSV are not packaged into virions, the encapsidation signal of PRRSV is likely localized in the ORF 1 region which is unique to the viral genome, but which is not present in the sg mRNAs.

In Experiment 2 below, the junction segments (the leader-mRNA junction sequences) of sg mRNAs 3 and 4 of two U.S. isolates of PRRSV, ISU 79 and ISU 1894, are determined. The knowledge of the leader-mRNA junction sequence identities provides means for effectively producing (a) chimeric viruses to be used as an infectious clone and/or as a vaccine, and (b)

vectors for inserting or "shuttling" one or more genes into a suitable, infectable host. Methods for designing and producing such chimeric viruses, infectious clones and vectors are known (see, for example, Sambrook et al, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

The leader-mRNA junction sequence of sg mRNAs 3 and 4 of the two isolates are different (TTGACC for mRNA 4-1 of ISU 79, GTAACC for mRNA 3, and TTCACC for mRNA 4). Most of the nucleotide differences in the junctions are present in the first 3 nucleotides. The last 3 nucleotides are invariable, suggesting that the joining of the leader sequence to the bodies of sg mRNAs occurs within the 5'-end of the leader-mRNA junction sequence. Similar observations have been reported for LV, EAV and LDV.

The acquisition of the additional sg mRNA 4-1 in isolate ISU 79 is due to a single nucleotide substitution which generates a new leader mRNA junction sequence. This substitution occurs in the last nucleotide of the junction segment, suggesting that the last nucleotide of the leader-mRNA junction motif is critical for the binding of the leader and for the initiation of transcription.

Although the sequence homology between the leader and the intergenic regions of coronaviruses led to the hypothesis that basepairing might be essential in the leader-primed transcription, no experimental evidence has documented for the

requirement of base-pairing in transcription of the sg mRNAs. For example, the sequence at the 3'-end of the leader of both coronaviruses and arteriviruses that is involved in the fusion process remains unknown.

Several lines of evidence support the leader-primed transcription mechanism for coronaviruses, but the presence of negative-stranded sg mRNAs and sg replicative intermediates (sg RI) in coronavirus-infected cells suggests that the mechanism involved in sg mRNA synthesis is more complex than mere base-pairing of the leader sequence with a junction sequence. However, negative-stranded sg mRNAs have not been detected in arteriviruses except for LDV, and sg RIs have been detected only in EAV-infected cells. Therefore, sg mRNA synthesis in arteriviruses, and particularly in PRRSV, may be less complicated than in coronaviruses.

Sequence analysis of the ORFs 2 to 7 of two U.S. PRRSV isolates and comparison of the sequences with LV reveals the heterogeneity of the leader-mRNA junction sequences. The presence of the leader-mRNA junction motifs at positions which do not correspond to a sg mRNA raises a question as to whether the short stretch of only six nucleotides which are conserved in the leader and junction sequences in the genomes of PRRSV and other arteriviruses is sufficient for efficient binding of the leader to these specific junction sites upstream of the ORFs. This apparent discrepancy, however, may be explained by the following two possibilities.

First, additional structural elements, such as secondary structures or the sequences surrounding the leader-mRNA junction segment, are expected to be involved in the fusion (binding) of the leader to the specific sites. It has been shown that, in MHV, the sequence flanking the consensus sequence (leader-mRNA junction sequence) of UCUAAAC affects the efficiency of sg DI RNA transcription, and that the consensus sequence was necessary but not sufficient in and of itself for the synthesis of the DI mRNA.

Second, the distance between two leader-mRNA junction regions may affect the transcription of sg mRNAs. It has been demonstrated that the downstream leader-mRNA junction region was suppressing sg DI RNA synthesis of MHV from the upstream leader-mRNA junction region. The suppression was significant when the two leader-mRNA junction sequence separation was less than 35 nucleotides. However, significant inhibition of larger sg DI RNA synthesis (from the upstream leader-mRNA junction sequence) was not observed when the two leader-mRNA junction regions were separated by more than 100 nucleotides.

The previously reported experimental results are consistent with the observations reported in Experiment 2 below, where an additional species of sg mRNA 4-1, in addition to the sg mRNA 4, is observed in some of the PRRSV isolates. The leader-mRNA junction sequences of sg mRNAs 4 and 4-1 in the Iowa strain of PRRSV are separated by about 226 nucleotides. Therefore, the synthesis of the larger sg mRNA

4-1 from the upstream leader-mRNA junction sequence is not suppressed by the presence of the downstream leader-mRNA 4 junction sequence.

In contrast, multiple potential leader-mRNA junction sequences were found at different positions upstream of ORFs 3, 5, 6 and 7, but there were no sg mRNAs corresponding to these leader-mRNA junction motifs in the Northern blot analysis. Most of these leader-mRNA junction sequences are separated by less than 50 nucleotides from the downstream leader-mRNA junction region, except for ORF 7 (in which the two potential leader-mRNA junction sequences are separated by 114 nucleotides). However, sg mRNA 7 in Northern blot analysis showed a widely-diffused band. Therefore, transcription of the larger sg mRNA 7 from the upstream leader-mRNA junction sequence may not be significantly suppressed by the downstream junction sequence, but it is not easily distinguishable from the abundant sg mRNA 7 by Northern blot analysis.

THE PRESENT POLYNUCLEOTIDES AND POLYPEPTIDES

ORF's 2-7 of plaque-purified PRRSV isolate ISU-12 (deposited on October 30, 1992, in the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., under the accession numbers VR 2385 [3 x plaque-purified] and VR 2386 [non-plaque-purified]) and ORF's 6-7 of PRRSV isolates ISU-22, ISU-55, ISU-3927 (deposited on

September 29, 1993, in the American Type Culture Collection under the accession numbers VR 2429, VR 2430 and VR 2431, respectively), ISU-79 and ISU-1894 (deposited on August 31, 1994, in the American Type Culture Collection under the accession numbers VR 2474 and VR 2475, respectively) are described in detail in U.S. Application Serial No. 08/301,435. However, the techniques used to isolate, clone and sequence these genes can be also applied to the isolation, cloning and sequencing of the genomic polynucleic acids of any PRRSV. Thus, the present invention is not limited to the specific sequences disclosed in the Experiments below.

For example, primers for making relatively large amounts of DNA by the polymerase chain reaction (and if desired, for making RNA by transcription and/or protein by translation in accordance with known *in vivo* or *in vitro* methods) can be designed on the basis of sequence information where more than one sequence obtained from a PRRSV genome has been determined (e.g., ORF's 2-7 of VR 2385, VR 2429, VR 2430, VR 2431, VR 2474, ISU-1894, VR 2332 and Lelystad virus). A region from about 15 to 50 nucleotides in length having at least 80% and preferably at least 90% identity is selected from the determined sequences. A region where a deletion occurs in one of the sequences (e.g., of at least 5 nucleotides) can be used as the basis for preparing a selective primer for selective amplification of the polynucleic acid of one strain or type of

PRRSV over another (e.g., for the differential diagnosis of North American and European PRRSV strains).

Once the genomic polynucleic acid is amplified and cloned into a suitable host by known methods, the clones can be screened with a probe designed on the basis of the sequence information disclosed herein. For example, a region of from about 50 to about 500 nucleotides in length is selected on the basis of either a high degree of identity (e.g., at least 90%) among two or more sequences (e.g., in ORF's 6-7 of the Iowa strains of PRRSV disclosed in Experiment III below), and a polynucleotide of suitable length and sequence identity can be prepared by known methods (such as automated synthesis, or restriction of a suitable fragment from a polynucleic acid containing the selected region, PCR amplification using primers which hybridize specifically to the polynucleotide, and isolation by electrophoresis). The polynucleotide may be labeled with, for example, ^{32}P (for radiometric identification) or biotin (for detection by fluorometry). The probe is then hybridized with the polynucleic acids of the clones and detected according to known methods.

The present Inventors have discovered that one or more of ORFs 2-4 may be related to the virulence of PRRSV. For example, at least one isolate of PRRSV which shows relatively low virulence also appears to have a deletion in ORF 4 (see, for example, Experiments VIII-XI in U.S. Application Serial No. 08/301,435). Furthermore, the least virulent known

isolate (VR 2431) shows a relatively high degree of variance in both nucleotide and amino acid sequence information in ORFs 2-4, as compared to other U.S. PRRSV isolates. Thus, in one embodiment, the present invention concerns polynucleotides and polypeptides related to ORFs 2-4 of VR 2431.

In a further embodiment, the present invention is concerned with a polynucleic acid obtained from a PRRSV isolate which confers immunogenic protection directly or indirectly against a subsequent challenge with a PRRSV, but in which the polynucleic acid is deleted or mutated to an extent which would render a PRRSV containing the polynucleic acid either low-virulent (i.e., a "low virulence" (lv) phenotype; see the corresponding explanation in U.S. Application Serial No. 08/301,435) or non-virulent (a so-called "deletion mutant"). Preferably, one or more of ORFs 2-4 is/are deleted or mutated to an extent which would render a PRRS virus non-virulent. However, it may be desirable to retain regions of one or more of ORFs 2-4 in the present polynucleic acid which (i) encode an antigenic and/or immunoprotective peptide fragment and which (ii) do not confer virulence to a PRRS virus containing the polynucleic acid.

The present invention also encompasses a PRRSV *per se* in which one or more of ORFs 2-4 is deleted or mutated to an extent which renders it either low-virulent or non-virulent (e.g., VR 2431). Such a virus is useful as a vaccine or as a vector for transforming a suitable host (e.g., MA-104, PSP 36,

CRL 11171, MARC-145 or porcine alveolar macrophage cells) with a heterologous gene. Preferred heterologous genes which may be expressed using the present deletion mutant may include those encoding a protein or an antigen other than a porcine reproductive and respiratory syndrome virus antigen (e.g., pseudorabies and/or swine influenza virus proteins and/or polypeptide-containing antigens, a porcine growth hormone, etc.) or a polypeptide-based adjuvant (such as those discussed in U.S. Application Serial No. 08/301,435 for a vaccine composition).

It may also be desirable in certain embodiments of the present polynucleic acid which contain, for example, the 3'-terminal region of a PRRSV ORF (e.g., from 200 to 700 nucleotides in length), at least part of which may overlap with the 5'-region of the ORF immediately downstream. Similarly, where the 3'-terminal region of an ORF may overlap with the 5'-terminal region of the immediate downstream ORF, it may be desirable to retain the 5'-region of the ORF which overlaps with the ORF immediately downstream.

The present Inventors have also discovered that ORF 5 in the PRRSV genome appears to be related to replication of the virus in mammalian host cells capable of sustaining a culture while infected with PRRSV. Accordingly, the present invention is also concerned with polynucleic acids obtained from a PRRSV genome in which ORF 5 may be present in multiple copies (a so-called "overproduction mutant"). For example, the present

polynucleic acid may contain at least two, and more preferably, from 2 to 10 copies of ORF 5 from a high-replication (hr) phenotype PRRSV isolate.

Interestingly, the PRRSV isolate ISU-12 has a surprisingly large number of potential start codons (ATG/AUG sequences) near the 5'-terminus of ORF 5, possibly indicating alternate start sites of this gene. Thus, alternate forms of the protein encoded by ORF 5 of a PRRSV isolate may exist, particularly where alternate ORF's encode a protein having a molecular weight similar to that determined experimentally (e.g., from about 150 to about 250 amino acids in length). The most likely coding region for ORF 5 of ISU-12 is indicated in Figure 1.

One can prepare deletion and overproduction mutants in accordance with known methods. For example, one can prepare a mutant polynucleic acid which contains a "silent" or degenerate change in the sequence of a region encoding a polypeptide. By selecting and making an appropriate degenerate mutation, one can substitute a polynucleic acid sequence recognized by a known restriction enzyme (see, for example, Experiment 2 below). Thus, if a silent, degenerate mutation is made at one or two of the 3'-end of an ORF and the 5'-end of a downstream ORF, one can insert a synthetic polynucleic acid (a so-called "cassette") which may contain a polynucleic acid encoding one or multiple copies of an hr ORF 5 protein product, of a PRRSV or other viral envelope protein

and/or an antigenic fragment of a PRRSV protein. The "cassette" may be preceded by a suitable initiation codon (ATG), and may be suitably terminated with a termination codon at the 3'-end (TAA, TAG or TGA). Of course, an oligonucleotide sequence which does not encode a polypeptide may be inserted, or alternatively, no cassette may be inserted. By doing so, one may provide a so-called deletion mutant.

The present invention also concerns regions and positions of the polypeptides encoded by ORFs of VR 2431 which may be responsible for the low virulence of this isolate. Accordingly, the present isolated and/or purified polypeptide may be one or more encoded by a "low-virulence mutation" of one or more of ORFs 2, 3 and 4 of a PRRSV (or a low-virulence fragment thereof at least 5 amino acids in length) in which one or more of positions 12-14 of the polypeptide encoded by ORF 2 are RGV (in which "R", "G" and "V" are the one-letter abbreviations for the corresponding amino acids), positions 44-46 are LPA, position 88 is A, position 92 is R, position 141 is G, position 183 is H, position 218 is S, position 240 is S and positions 252-256 are PSSSW, or any combination thereof. Other amino acid residue identities which can be further combined with one or more of the above amino acid position identities include those at position 174 (I) and position 235 (M).

The present isolated and/or purified polypeptide may also be one encoded by an ORF 3 of a PRRSV in which one or more of the specified amino acid identities may be selected from those at positions 11 (L), 23 (V), 26-28 (TDA), 65-66 (QI), 70 (N), 79 (N), 93 (T), 100-102 (KEV), 134 (K), 140 (N), 223-227 (RQRIS), 234 (A) and 235 (M), or any combination thereof, which may be further combined with one or more of positions 32 (F), 38 (M), 96 (P), 143 (L), 213-217 (FQTS), 231 (R), and 252 (A).

The present isolated and/or purified polypeptide may also be one encoded by an ORF 4 of a PRRSV in which one or more of the specified amino acid identities may be selected from those at positions 13 (E), 43 (N), 56 (G), 58-59 (TT), 134 (T), 139 (I) and any combination thereof, which may be further combined with one or more of positions 2-3 (AA), 51 (G) and 63 (P).

The present invention also concerns polynucleotide sequences encoding polypeptide sequences of 5 or more amino acids, preferably 10 or more amino acids, and up to the full length of the polypeptide, encoded by any one of ORFs 2-4 of VR 2431, in which the polynucleotides at the codon(s) corresponding to the amino acid positions detailed in the preceding three paragraphs are replaced with polynucleotides encoding the corresponding amino acids of the proteins encoded by the corresponding ORF of VR 2431.

In a further embodiment of the present invention, the polynucleic acid encodes one or more proteins, or antigenic

regions thereof, of a PRRSV. Preferably, the present nucleic acid encodes at least one antigenic region of a PRRSV membrane (envelope) protein. More preferably, the present polynucleic acid encodes a hypervariable region from a ORF 5 PRRSV protein product (see the discussion below) or (b) contains at least one copy of the ORF-5 gene from a high virulence (hv) phenotype isolate of PRRSV (see the description of "hv phenotype" in U.S. Application Serial No. 08/301,435) and a sufficiently long fragment, region or sequence of at least one of ORF-2, ORF-3, ORF-4, ORF-5 and/or ORF-6 from the genome of a PRRSV isolate to encode an antigenic region of the corresponding protein(s) and effectively stimulate protection against a subsequent challenge with, for example, a hv phenotype PRRSV isolate.

Even more preferably, at least one entire envelope protein encoded by ORF-2, ORF-3, ORF-5 and/or ORF-6 of a PRRSV is contained in the present polynucleic acid, and the present polynucleic acid excludes or modifies a sufficiently long portion of one of ORFs 2-4 from a PRRSV to render a PRRSV containing the same either low-virulent or non-virulent. Most preferably, the polynucleic acid is isolated from the genome of an isolate of the Iowa strain of PRRSV (for example, VR 2385 (3X plaque-purified ISU-12), VR 2386 (non-plaque-purified ISU-12), VR 2428 (ISU-51), VR 2429 (ISU-22), VR 2430 (ISU-55), VR 2431 (ISU-3927), VR 2474 (ISU-79) and/or ISU-1894).

A further preferred embodiment of the present invention includes a polynucleotide encoding an amino acid sequence from a hypervariable region of ORF 5 of a PRRSV, preferably of an Iowa strain of PRRSV. Thus, such polynucleotides encode one (or more) of the following amino acid sequences:

TABLE 1

Hypervariable Region 1 (positions 32-38)	Hypervariable Region 2 (Positions 57-66)	Hypervariable Region 3 (Pos'ns 120-128)
NGNSGSN	ANKFDWAVET	LICFVIRLA
SNDSSSH	ANKFDWAVEP	LTCFVIRFA
SSSNSSH	AGEFDWAVET	LICFVIRFT
SANSSSH	ADKFDWAVEP	LACFVIRFA
HSNSSSH	ADRFDWAVEP	LTCFVIRFV
SNSSSSH	SSHFGWAVET	LTCFIIRFA
NNSSSSH		FICFVIRFA
NGGDSST(Y)		FVCFVIRAA

In this embodiment, the polynucleotide may encode further amino acid sequences of a PRRSV ORF 5 (as disclosed in Figure 3 or in U.S. Application Serial Nos. 08/131,625 or 08/301,435), as long as one or more of the hypervariable regions at positions 32-38, 57-66 and/or 120-128 are included. (The present invention specifically excludes the proteins and polynucleotides of ORF 5 of LV and VR 2332.)

A further preferred embodiment of the present invention concerns a purified preparation which may comprise, consist essentially of or consist of a polynucleic acid having a sequence of the formula (I) or (II):

$$5'-\alpha-\beta-3' \qquad \text{(I)}$$

$$5'-\alpha-\beta-\gamma-3' \qquad \text{(II)}$$

wherein α encodes at least one polypeptide, or antigenic or low-virulence fragment thereof encoded by a polynucleotide selected from the group consisting of ORFs 2, 3 and 4 of an Iowa strain of PRRSV and regions thereof encoding such antigenic and/or low-virulence fragments; and β is at least one copy of an ORF 5 from an Iowa strain of PRRSV or an antigenic fragment thereof (e.g. one or more hypervariable regions), preferably a full-length copy from a high replication (hr) phenotype; and γ encodes at least one polypeptide or antigenic fragment thereof encoded by a polynucleotide selected from the group consisting of ORF 6 and ORF 7 of an Iowa strain of PRRSV and regions thereof encoding the antigenic fragments.

Alternatively, the present invention may concern a purified preparation which may comprise, consist essentially of or consist of a polynucleic acid having a sequence of the formula (III):

$$5'-\beta-\delta-\gamma-3' \qquad \text{(III)}$$

where β and γ are as defined above; and δ is either a covalent bond or a linking polynucleic acid which does not materially affect transcription and/or translation of the polynucleic

acid. Preferably, β is a polynucleotide encoding at least one hypervariable region of a protein encoded by an ORF 5 of an Iowa strain of PRRSV, and more preferably, encodes a full-length protein encoded by an ORF 5 of an Iowa strain of PRRSV.

The present invention may also concern a purified preparation which may comprise, consist essentially of or consist of a polynucleic acid having a sequence of the formula (IV):

$$5'-\alpha-\beta-\delta-\gamma-3' \qquad \text{(IV)}$$

where α, β, γ and δ are as defined in formulas (I)-(III) above.

The present invention may also concern a purified preparation which may comprise, consist essentially of or consist of a polynucleic acid, an expression vector or a plasmid having a sequence of the formula (V):

$$5'-\epsilon-\zeta-\iota-\kappa-\xi-3' \qquad \text{(V)}$$

where ϵ, which is optionally present, is a 5'-terminal polynucleotide sequence which provides a means for operationally expressing the polynucleotides α, β, γ and δ; ζ is a polynucleotide of the formula KTVACC, where K is T, G or U, and V is A, G or C; ι is a polynucleotide of at most about 130 (preferably at most 100) nucleotides in length; κ is a

polynucleotide comprising one or more genes selected from the group consisting of a conventional marker or reporter gene, α, β, γ, and operationally linked combinations thereof, where α, β, and γ are as defined in formulas (I)-(IV) above; and ξ, which is optionally present, is a 3'-terminal polynucleotide sequence which does not suppress the operational expression of the polynucleotides α, β, γ and δ, and which may be operationally linked to ϵ (for example, in a plasmid).

Suitable marker or reporter genes include, e.g., those providing resistance to an antibiotic such as neomycin, erythromycin or chloramphenicol; those encoding a known, detectable enzyme such as β-lactamase, DHFR, horseradish peroxidase, glucose-6-phosphate dehydrogenase, alkaline phosphatase, and enzymes disclosed in U.S. Patent 4,190,496, col. 32, line 33 through col. 38, line 44 (incorporated herein by reference), etc.; and those encoding a known antibody (e.g., mouse IgG, rabbit IgG, rat IgG, etc.) or known antigenic protein such as Protein A, Protein G, bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), bovine gamma globulin (BGG), lactalbumin, polylysine, polyglutamate, lectin, etc.

The polynucleotide ι is preferably a polynucleotide sequence at least 80% homologous to a polynucleotide sequence from a PRRSV genome located between a leader-mRNA junction sequence and the start codon of the ORF immediately downstream. "About 130" nucleotides in length refers to a

length of the polynucleotide ι which does not adversely affect the operational expression of κ. For example, in ISU 79, a leader-mRNA junction sequence which does not suppress expression of ORF 7 can be found 129 bases upstream from the start codon of ORF 7 (see Experiment 2 below). Suitable exemplary sequences for the polynucleotide ι can be deduced from the sequences shown in Figs. 1 and 9.

The present polynucleic acid may also comprise, consist essentially of or consist of combinations of the above sequences, either as a mixture of polynucleotides or covalently linked in either a head-to-tail (sense-antisense) or head-to-head fashion. Polynucleic acids complementary to the above sequences and combinations thereof (antisense polynucleic acid) are also encompassed by the present invention. Thus, in addition to possessing multiple or variant copies of ORF 5, the present polynucleic acid may also contain multiple or variant copies of one or more of ORF's 1-7, including antigenic or hypervariable regions of ORF 5, of Iowa strain PRRSV's.

Similar to the methods described above and in the Experiments described below and in U.S. Application Serial Nos. 08/131,625 and 08/301,435, one can prepare a library of recombinant clones (e.g., using *E. coli* as a host) containing suitably prepared restriction fragments of a PRRSV genome (e.g., inserted into an appropriate plasmid expressible in the host). The clones are then screened with a suitable probe

(e.g, based on a conserved sequence of ORF's 2-3; see, for example, Figure 22 of U.S. Application Serial No. 08/301,435). Positive clones can then be selected and grown to an appropriate level. The polynucleic acids can then be isolated from the positive clones in accordance with known methods. A suitable primer for PCR can then be designed and prepared as described above to amplify the desired region of the polynucleic acid. The amplified polynucleic acid can then be isolated and sequenced by known methods.

The present purified preparation may also contain a polynucleic acid selected from the group consisting of sequences having at least 97% sequence identity (or homology) with at least one of ORFs 5-7 of VR 2385, VR 2430 and/or VR 2431; and sequences encoding a polypeptide having at least the minimum sequence identity (or homology) with at least one of ORF's 2-5 of VR 2385, VR 2428, VR 2429, VR 2430, VR 2431, VR 2474 and ISU-1894, as follows:

TABLE 2

Relative to Isolate:	Minimum % Homology with ORF:			
	2	3	4	5
VR 2385	99	92	95	90
VR 2429	100	99	99	98
VR 2430	98	95	96	90
VR 2431	94	88	93	92
VR 2474	99	97	97	95
ISU 1894	97	97	99	97

Preferably, the polynucleic acid excludes or modifies a sufficiently long region or portion of one or more of ORFs 2-4 of the hv PRRSV isolates VR 2385, VR 2429, ISU-28, ISU-79 and/or ISU-984 to render the isolate low-virulent or non-virulent.

In the context of the present application, "homology" refers to the percentage of identical nucleotide or amino acid residues in the sequences of two or more viruses, aligned in accordance with a conventional method for determining homology (e.g., the MACVECTOR or GENEWORKS computer programs, aligned in accordance with the procedure described in Experiment III in U.S. Application Serial No. 08/301,435).

Preferably, the present isolated polynucleic acid encodes a protein, polypeptide, or antigenic fragment thereof which is at least 10 amino acids in length and in which non-homologous amino acids which are non-essential for antigenicity may be conservatively substituted. An amino acid residue in a protein, polypeptide, or antigenic fragment thereof is conservatively substituted if it is replaced with a member of its polarity group as defined below:

Basic amino acids:

lysine (Lys), arginine (Arg), histidine (His)

Acidic amino acids:

aspartic acid (Asp), glutamic acid (Glu), asparagine (Asn), glutamine (Gln)

Hydrophilic, nonionic amino acids:

serine (Ser), threonine (Thr), cysteine (Cys), asparagine (Asn), glutamine (Gln)

Sulfur-containing amino acids:

cysteine (Cys), methionine (Met)

Hydrophobic, aromatic amino acids:

phenylalanine (Phe), tyrosine (Tyr), tryptophan (Trp)

Hydrophobic, nonaromatic amino acids:

glycine (Gly), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), proline (Pro)

More particularly, the present polynucleic acid encodes one or more of the protein(s) encoded by the second, third, fourth, fifth, sixth and/or seventh open reading frames (ORF's 2-7) of the PRRSV isolates VR 2385, VR 2386, VR 2428, VR 2429, VR 2430, VR 2431, VR 2474 and/or ISU-1894 (e.g., one or more of the sequences shown in Fig. 3 and/or SEQ ID NOS:15, 17, 19, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63 and 65 of U.S. Application Serial No. 08/301,435).

ORF's 6 and 7 are not likely candidates for controlling virulence and replication phenotypes of PRRSV, as the nucleotide sequences of these genes are highly conserved among high virulence (hv) and low virulence (lv) isolates (see Experiment III of U.S. Application Serial No. 08/301,435). However, ORF 5 in PRRSV isolates appears to be less conserved among high replication (hr) and low replication (lr) isolates. Therefore, it is believed that the presence of an ORF 5 from an hr PRRSV isolate in the present polynucleic acid will enhance the production and expression of a recombinant vaccine produced from the polynucleic acid.

Furthermore, ORF 5 of PRRSV contains three hydrophilic, hypervariable regions typically associated with antigenicity in a polypeptide. Thus, the present invention also

encompasses polynucleotides encoding a polypeptide comprising one or more hypervariable regions of a PRRSV ORF 5, preferably a polypeptide of the formula *a-b-c-d-e-f-g*, where:

a is an amino group, a poly(amino acid) corresponding to positions 1-31 of a protein encoded by a PRSSV ORF 5, or a fragment of such a poly(amino acid) which does not adversely affect the antigenicity of the polypeptide;

b is an amino acid sequence selected from the group consisting of those sequences listed under Hypervariable Region No. 1 in Table 1 above,

c is an amino acid sequence corresponding to positions 39-56 of a protein encoded by a PRSSV ORF 5 (preferably a sequence of the formula LQLIYNLTLCELNGTDWL, in which one or more [preferably 1-10] amino acids may be conservatively substituted),

d is an amino acid sequence selected from the group consisting of those sequences listed under Hypervariable Region No. 2 in Table 1 above,

e is an amino acid sequence corresponding to positions 67-119 of a protein encoded by a PRRSV ORF 5, in which one or more (preferably 1-20, and more preferably 1-10) amino acid residues may be conservatively substituted and which does not adversely affect the antigenicity of the polypeptide,

f is an amino acid sequence selected from the group consisting of those sequences listed under Hypervariable Region No. 3 in the Table above, and

g is a carboxy group (a group of the formula -COOH), an amino acid sequence corresponding to positions 129-200 of a protein encoded by a PRSSV ORF 5 or a fragment thereof which does not adversely affect the antigenicity of the polypeptide.

Accordingly, it is preferred that the present polynucleic acid, when used for immunoprotective purposes (e.g., in the preparation of a vaccine), contain at least one copy of ORF 5 from a high-replication isolate (i.e., an isolate which grows to a titer of 10^6-10^7 $TCID_{50}$ in, for example, CRL 11171 cells; also see the discussions in Experiments VIII-XI U.S. Application Serial No. 08/301,435).

On the other hand, the lv isolate VR 2431 appears to be a deletion mutant, relative to hv isolates (see Experiments III and VIII-XI U.S. Application Serial No. 08/301,435). The deletion appears to be in ORF 4, based on Northern blot analysis. Accordingly, when used for immunoprotective purposes, the present polynucleic acid preferably does not contain a region of ORF 4 from an hv isolate responsible for high virulence, and more preferably, excludes the region of ORF 4 which does not overlap with the adjacent ORF's 3 and 5.

It is also known (at least for PRRSV) that neither the nucleocapsid protein nor antibodies thereto confer immunological protection against PRRSV to pigs. Accordingly, the present polynucleic acid, when used for immunoprotective purposes, contains one or more copies of one or more regions from ORF's 2, 3, 4, 5 and 6 of a PRRSV isolate encoding an

antigenic region of the viral envelope protein, but which does not result in the symptoms or histopathological changes associated with PRRS when administered to a pig. Preferably, this region is immunologically cross-reactive with antibodies to envelope proteins of other PRRSV isolates.

Similarly, the protein encoded by the present polynucleic acid confers protection against PRRS to a pig administered a composition comprising the protein, and antibodies to this protein are immunologically cross-reactive with the envelope proteins of other PRRSV isolates. More preferably, the present polynucleic acid encodes the entire envelope protein of a PRRSV isolate or a protein at least 80% homologous thereto and in which non-homologous residues are conservatively substituted, or alternatively a protein at least 98% homologous thereto. Most preferably, the present polynucleotide is one of the sequences shown in Fig. 1, encompassing at least one of the open reading frames recited therein.

Relatively short segments of polynucleic acid (about 20 bp or longer) in the genome of a virus can be used to screen or identify tissue and/or biological fluid samples from infected animals, and/or to identify related viruses, by methods described herein and known to those of ordinary skill in the fields of veterinary and viral diagnostics and veterinary medicine. Accordingly, a further aspect of the present invention encompasses an isolated (and if desired,

purified) polynucleic acid consisting essentially of a fragment of from 15 to 2000 bp, preferably from 18 to 1000 bp, and more preferably from 21 to 100 bp in length, derived from ORF's 2-7 of a PRRSV genome (preferably the Iowa strain of PRRSV). Particularly preferably, the present isolated polynucleic acid fragments are obtained from a terminus of one or more of ORF's 2-7 of the genome of the Iowa strain of PRRSV, and most preferably, are selected from the group consisting of the primers described in Experiments 1 and 2 below and SEQ ID NOS:1-12, 22 and 28-34 of U.S. Application Serial No. 08/301,435.

The present invention also concerns a diagnostic kit for assaying a porcine reproductive and respiratory syndrome virus, comprising (a) a first primer comprising a polynucleotide having a sequence of from 10 to 50 nucleotides in length which hybridizes to a genomic polynucleic acid from an Iowa strain of porcine reproductive and respiratory syndrome virus at a temperature of from 25 to 75°C, (b) a second primer comprising a polynucleotide having a sequence of from 10 to 50 nucleotides in length, said sequence of said second primer being found in said genomic polynucleic acid from said Iowa strain of porcine reproductive and respiratory syndrome virus and being downstream from the sequence to which the first primer hybridizes, and (c) a reagent which enables detection of an amplified polynucleic acid. Preferably, the

reagent is an intercalating dye, the fluorescent properties of which change upon intercalation into double-stranded DNA.

The present isolated polynucleic acid fragments can be obtained by: (i) digestion of the cDNA corresponding to (complementary to) the viral polynucleic acids with one or more appropriate restriction enzymes, (ii) amplification by PCR (using appropriate primers complimentary to the 5′ and 3′-terminal regions of the desired ORF(s) or to regions upstream of the 5′-terminus or downstream from the 3′-terminus) and cloning, or (iii) synthesis using a commercially available automated polynucleotide synthesizer.

Another embodiment of the present invention concerns one or more proteins or antigenic fragments thereof from a PRRS virus, preferably from the Iowa strain of PRRSV. As described above, an antigenic fragment of a protein from a PRRS virus (preferably from the Iowa strain of PRRSV) is at least 5 amino acids in length, particularly preferably at least 10 amino acids in length, and provides or stimulates an immunologically protective response in a pig administered a composition containing the antigenic fragment.

Methods of determining the antigenic portion of a protein are known to those of ordinary skill in the art (see the description above). In addition, one may also determine an essential antigenic fragment of a protein by first showing that the full-length protein is antigenic in a host animal (e.g., a pig). If the protein is still antigenic in the

presence of an antibody which specifically binds to a particular region or sequence of the protein, then that region or sequence may be non-essential for immunoprotection. On the other hand, if the protein is no longer antigenic in the presence of an antibody which specifically binds to a particular region or sequence of the protein, then that region or sequence is considered to be essential for antigenicity.

Three hypervariable regions in ORF 5 of PRRSV have been identified by comparing the amino acid sequences of the ORF 5 product of all available PRRSV isolates (see, for example, Fig. 2D). Amino acid variations in these three regions are significant, and are not structurally conserved (Fig. 2D). All three hypervariable regions are hydrophilic and antigenic. Thus, these regions are likely to be exposed to the viral membrane and thus be under host immune selection pressure. hypervariable regions as antigenic determinants in the ORF 5 envelope protein.

The present invention also concerns a protein or antigenic fragment thereof encoded by one or more of the polynucleic acids defined above, and preferably by one or more of the ORF's of a PRRSV, more preferably of the Iowa strain of PRRSV. The present proteins and antigenic fragments are useful in immunizing pigs against PRRSV, in serological tests for screening pigs for exposure to or infection by PRRSV (particularly the Iowa strain of PRRSV), etc.

For example, the present protein may be selected from the group consisting of the proteins encoded by ORF's 2-7 of VR 2385, ISU-22 (VR 2429), ISU-55 (VR 2430), ISU-1894, ISU-79 (VR 2474) and ISU-3927 (VR 2431) (e.g., one or more of the sequences shown in Fig. 2 and/or SEQ ID NOS:15, 17, 19, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 67, 69 and 71 of U.S. Application Serial No. 08/301,435); antigenic regions of at least one of these proteins having a length of from 5 amino acids to less than the full length of the protein; polypeptides having the minimum homology with the protein encoded by the PRSSV ORF indicated in Table 2 above; and polypeptides at least 97% homologous with a protein encoded by one of the ORF's 6-7 of VR 2385, VR 2429, VR 2430, ISU-1894, ISU-79 and VR 2431 (e.g., SEQ ID NOS:17, 19, 43, 45, 47, 49, 51, 53, 55, 57, 59 and 61 of U.S. Application Serial No. 08/301,435). Preferably, the present protein has a sequence encoded by an ORF selected from the group consisting of ORFs 2-5 of VR 2385, VR 2428, VR 2429, VR 2430, VR 2431, VR 2474 and ISU-1894 (see, for example, Fig. 2A-D); variants thereof which provide effective immunological protection to a pig administered the same and in which from 1 to 100 (preferably from 1 to 50 and more preferably from 1 to 25) deletions or conservative substitutions in the amino acid sequence exist; and antigenic fragments thereof at least 5 and preferably at least 10 amino acids in length which provide effective immunological protection to a pig administered the same.

More preferably, the present protein variant or protein fragment has a binding affinity (or association constant) of at least 1% and preferably at least 10% of the binding affinity of the corresponding full-length, naturally-occurring protein to a monoclonal antibody which specifically binds to the full-length, naturally-occurring protein (i.e., the protein encoded by a PRRSV ORF).

The present invention also concerns a method of producing a polypeptide, comprising expressing the present polynucleic acid in an operational expression system, and purifying the expressed polypeptide from the expression system. Suitable expression systems include those conventionally used for either *in vitro* or *in vivo* expression of proteins and polypeptides, such as a rabbit reticulocyte system for *in vitro* expression, and for *in vivo* expression, a modified or chimeric PRRSV (used to infect an infectable host cell line, such as MA-104, CRL 11171, PSP-36, PSP-36-SAH, MARC-145 and porcine alveolar macrophages), or a conventional expression vector containing the present polynucleic acid, under the operational control of a known promoter (e.g., a thymidine kinase promoter, SV40, etc.) for use in conventional expression systems (e.g., bacterial plasmids and corresponding host bacteria, yeast expression systems and corresponding host yeasts, etc.). The expressed polypeptide or protein is then purified or isolated from the expression system by conventional purification and/or isolation methods.

Other features of the invention will become apparent in the course of the following descriptions of exemplary embodiments, which are given for illustration of the invention, and are not intended to be limiting thereof.

EXPERIMENT I

Summary:

The sequences of ORFs 2 to 5 of one low virulence, one "moderate" virulence and one high virulence U.S. PRRSV isolate have been determined and analyzed. Comparisons with known sequences of other PRRSV isolates show that considerable sequence variations at both nucleotide and amino acid levels exist in ORFs 2 to 5 of seven U.S. isolates with differing virulence. However, ORFs 6 and 7 of these seven U.S. isolates are highly conserved (U.S. Application Serial No. 08/301,435). Extensive sequence variations were also found in ORFs 2 to 7 between the European LV and the U.S. isolates. The least virulent U.S. PRRSV isolate known (ISU-3927) displayed the most sequence variation, in comparison with other U.S. isolates.

The phylogenetic relationship of the U.S. isolates was also analyzed. Phylogenetic analysis of the ORFs 2 to 7 of the U.S. isolates indicated that there are at least three groups of PRRSV variants (or minor genotypes) within the major U.S. PRRSV genotype. Consequently, it is highly likely that a number of additional major or minor genotypes will be

identified as more virus isolates from different geographic regions are examined.

Interestingly, the least virulent U.S. isolate known (ISU 3927) forms a branch distinct from other U.S. isolates. Analysis of the nucleotide and amino acid sequences also showed that the isolate ISU 3927 exhibits the most variations in ORFs 2 to 4, relative to other U.S. isolates. Many of these variations in isolate ISU 3927 result in non-conserved amino acid substitutions. However, these non-conserved changes in isolate ISU 3927, as compared to other U.S. isolates, do not appear to be limited to a particular region; they are present throughout ORFs 2 to 4. Therefore, a specific correlation between sequence variations and viral virulence is not yet fully elucidated (although certain positions in ORF 3 appear to be possibly related to virulence; see Fig. 2B, positions 30, 48, 54-56, 134, 140, 143, 147, 153, 206, and 215; amino acids at one or more of these positions may serve as a basis for mutating other known proteins encoded by a PRRSV ORF 3).

Results:

The amino acid sequence identity between seven U.S. PRRSV isolates was 91-99% in ORF 2, 86-98% in ORF 3, 92-99% in ORF 4 and 88-97% in ORF 5. The least virulent U.S. isolate known has higher sequence variations in the ORFs 2 to 4 than in ORFs 5 to 7, as compared to other U.S. isolates. Three

hypervariable regions with antigenic potential were identified in the major envelope glycoprotein encoded by ORF 5.

Pairwise comparison of the sequences of ORFs 2 to 7 and phylogenetic tree analysis implied the existence of at least three groups of PRRSV variants (or minor genotypes) within the major genotype of U.S. PRRSV. The least virulent U.S. isolate known forms a distinct branch from other U.S. isolates with differing virulence. The results of this study have implications for the taxonomy of PRRSV and vaccine development.

Figure 1 shows a nucleotide sequence comparison of ORFs 2 to 5 of U.S. isolates ISU 3927, ISU 22 and ISU 55 with other known PRRSV isolates. The nucleotide sequence of VR 2385 is shown on top, and only differences are indicated. The start codon of each ORF is indicated by +>, and the termination codon of each ORF is indicated by asterisks (*). The leader-mRNA junction sequences for subgenomic mRNAs 3, 4 and 4-1 are underlined, and the locations of the junction sequences relative to the start codon of each ORF are indicated by minus (-) numbers of nucleotides upstream of each ORF. The sequences of VR 2385 (U.S. Application Serial Nos. 08/131,625 and 08/301,435), VR 2332, ISU 79 and ISU 1894 (U.S. Application Serial No. 08/301,435) used in this alignment were previously reported.

MATERIALS AND METHODS:

Cells and viruses:

The ATCC CRL 11171 cell line was used to propagate the PRRSV. The cells were grown in Dulbecco's minimal essential medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1 X antibiotics (penicillin G 10,000 unit/ml, streptomycin 10,000 mg/ml and amphotericin B 25 mg/ml).

Three U.S. isolates of PRRSV used in this study, designated as ISU 22, ISU 55 and ISU 3927, were isolated from pig lungs obtained from different farms in Iowa during PRRS outbreaks. All three isolates were plaque-purified three times on CRL 11171 cells before further experimentation. Comparative pathogenicity studies showed that isolate ISU 3927 is the least virulent isolate among 10 different U.S. PRRSV isolates. Isolate ISU 22 is a high virulence isolate and isolate ISU 55 is "moderately" pathogenic. All of the three virus isolates used in this experiment were at seventh passage.

Isolation of PRRSV intracellular RNAs:

Confluent monolayers of CRL 11171 cells were infected with the three U.S. isolates of PRRSV, ISU 22, ISU 55 and ISU 3927, respectively, at a multiplicity of infection (m.o.i.) of 0.1. At 24 hrs. postinfection, the infected cells were washed three times with cold PBS buffer. The total intracellular RNAs were then isolated by guanidinium isothiocyanate and

phenol-chloroform extraction (Stratagene). The presence of virus-specific RNA species in the RNA preparation was confirmed by Northern blot hybridization (data not shown). The total intracellular RNAs were quantified spectrophotometrically.

Reverse transcription and polymerase chain reaction (RT-PCR):

First strand complementary (c) DNA was synthesized from the total intracellular RNAs by reverse transcription using random primers as described previously (Meng et al., 1993, J. Vet. Diagn. Invest., 5:254-258). For amplification of the entire protein coding regions of the ORFs 2 to 5 of the three isolates of PRRSV, two sets of primers were designed on the basis of the sequences of VR 2385 and LV. Primers JM259 (5'-GGGGATCCTTTTGTGGAGCCGT-3') and JM260 (5'-GGGGAATTCGGGATAGGGAATGTG-3') amplified the sequence of ORFs 4 and 5, and primers XM992 (5'-GGGGGATCCTGTTGG-TAATAG(A)GTCTG-31 and XM993 (5'-GGTGAATTCGTTTTATTTCCCTCCGGGC-3') amplified the sequence of ORFs 2 and 3. Unique restriction sites (EcoRI or BamHI) at the 5' end of these primers were introduced to facilitate cloning. A degenerate base, G (A), was synthesized in primer XM 992 based on the sequences of VR 2385 and LV (Meulenberg et al., 1993; U.S. Application Serial No. 08/301,435). PCR was performed as described previously (Meng et al., 1993, J. Vet. Diagn. Invest., 5:254-258).

Cloning and nucleotide sequencing:

The RT-PCR products were analyzed by a 0.8% agarose gel electrophoresis. The two PCR fragments representing ORFs 2 and 3 as well as ORFs 4 and 5, respectively, were purified by the glassmilk procedure (GENECLEAN kit, BIO 101, Inc.). The purified fragments were each digested with BamHI and EcoRI, and cloned into the vector pSK+ as described previously (Meng et al., 1993). The E. *Coli* DH 5α cells were used for transformation of recombinant plasmids. White colonies were selected and grown in LB broth containing 100 mg/ml ampicillin. The *E. Coli* cells containing recombinant plasmid were lysed with lysozyme, and the plasmids were then isolated by using the Qiagen column (QIAGEN Inc.).

Plasmids containing viral inserts were sequenced with an automated DNA Sequencer (Applied Biosystem, Inc.). Three or more independent CDNA clones representing the entire sequence of ORFs 2 to 5 from each of the three PRRSV isolates were sequenced with universal and reverse primers. Several virus-specific primers, XM969 (5'-GATAGAGTCTGCCCTTAG-3'), XM970 (5'-GGTTTCACCTAGAATGGC-3'), XM1006 (5'-GCTTCTGAGATGAGTGA-3'), XM077 (5'-CAACCAGGCGTAAACACT-3') and XM078 (5'-CTGAGCAATT ACAGAAG-3'), were also used to determine the sequence of ORFs 2 to 5.

Sequence analyses:

Sequence data were combined and analyzed by using MacVector (International Biotechnologies, Inc.) and GeneWorks (IntelliGenetics, Inc.) computer software programs. Phylogenetic analyses were performed using the PAUP software package version 3.1.1 (David L. Swofford, Illinois Natural History Survey, Champaign, IL). PAUP employs the maximum parsimony algorithm to construct phylogenetic trees.

RESULTS:

Nucleotide sequence analyses of ORFs 2 to 5:

The sequences of ORFs 2 to 5 of five PRRSV isolates, ISU 79, ISU 1894, ISU 22, ISU 55 and ISU 3927, were determined and compared with other known PRRSV isolates including VR 2385, VR 2332 and LV (Meulenberg et al., 1993). The sequences of ORFs 6 and 7 of isolates VR 2385, ISU 22, ISU 55, ISU 79, ISU 1894 and ISU 3927 were reported previously (U.S. Application Serial No. 08/301,435). The isolates used in this experiment have been shown to differ in pneumovirulence in experimentally-infected pigs (U.S. Application Serial Nos. 08/131,625 and 08/301,435). ISU 3927 is the least virulent isolate among ten different U.S. PRRSV isolates (U.S. Application Serial No. 08/131,625 and U.S. Application Serial No. 08/301,435).

Like other U.S. PRRSV isolates, ORFs 2 to 4 of these isolates overlapped each other (Fig. 1). However, unlike LV, ORFs 4 and 5 of the U.S. isolates are separated by 10

nucleotides (Fig. 1). ORFs 4 and 5 of LV overlapped by one nucleotide. The single nucleotide substitution from A of the start codon of ORF 5 in LV to T in the U.S. isolates places the start codon of ORF 5 of the U.S. isolates 10 nucleotides downstream of the ORF 4 stop codon. Therefore, a 10-nucleotide noncoding sequence appears between ORFs 4 and 5 of the known U.S. isolates (Fig. 1).

ORF 2 of ISU 79 is 3 nucleotides shorter than other U.S. isolates. The single nucleotide substitution from TGG to TAG just before the stop codon of ORF 2 creates a new stop codon in ISU 79 (Fig. 1). A 3-nucleotide deletion was also found in ORF 5 of ISU 3927, compared to other U.S. isolates (Fig. 1). The size of ORFs 2 to 5 of all the U.S. isolates are identical, except for the ORF 2 of ISU 79 and ORF 5 of ISU 3927, both of which are 3 nucleotides shorter than the other ORFs (Fig. 1).

Sequence comparisons of ORFs 2 to 5 of the seven U.S. PRRSV isolates shown in Fig. 1 indicate that there are considerable nucleotide sequence variations in ORFs 2 to 5 of the U.S. isolates (Fig. 1). The nucleotide sequence identity was 96-98% in ORF 2, 92-98% in ORF 3, 92-99% in ORF 4, and 90-98% in ORF 5 between VR 2385, VR 2332, ISU 22, ISU 55, ISU 79, and ISU 1894 (Table 3).

The least virulent isolate ISU 3927 has the most variations among the seven U.S. isolates (Fig. 1 and Table 3). The nucleotide sequence identity between ISU 3927 and other

U.S. isolates was 93-94% in ORF 2, 89-90% in ORF 3, and 91-93% in ORF 4 (Table 3). Like ORFs 6 and 7 (U.S. Application Serial No. 08/301,435), ORF 5 of ISU 3927 has no significant changes except for a 3-nucleotide deletion (Fig. 1). ORF 5 of ISU 3927 shares 91-93% nucleotide sequence identity with the ORF 5 of other U.S. isolates (Table 3).

However, extensive sequence variation was found in ORFs 2 to 5 between LV and the U.S. isolates (Fig. 1 and Table 3). The nucleotide sequence identity between LV and the U.S. isolates was 65-67% in ORF 2, 61-64% in ORF 3, 63-66% in ORF 4, and 61 -63% in ORF 5 (Table 3). Extensive genetic variations in ORFs 6 and 7 between LV and U.S. PRRSV also exists (U.S. Application Serial Nos. 08/131,625 and 08/301,435). These results indicate that the least virulent isolate ISU 3927 is also the most distantly related of the U.S. isolates, with genetic variations occurring mostly in ORFs 2 to 4.

The single nucleotide substitution from TGG to TAG before the stop codon in ORF 2 observed in ISU 79 was also present in isolates ISU 55 and ISU 3927, both of which produce seven sg mRNAs, but not in isolates ISU 22, ISU 1894 or VR 2385, which each synthesize only six sg mRNAs (U.S. Application Serial Nos. 08/131,625 and 08/301,435). The results indicate that the leader-mRNA 4-1 junction sequence of ISU 55 and ISU 3927 is very likely to be the same as ISU 79 (Fig. 1).

The leader-mRNA junction sequences for sg mRNAs 3 and 4 of ISU 79 and ISU 1894 were determined to be GUAACC at 89 nucleotides upstream of ORF 3 for sg mRNA 3, and UUCACC at 10 nucleotides upstream of ORF 4 for sg mRNA 4 (U.S. Application Serial No. 08/301,435; see also Experiment 2 below). A sequence comparison of isolates ISU 22, ISU 55 and ISU 3927 with isolates VR 2385, ISU 79 and ISU 1894 indicates that the leader-mRNA junction sequences for sg mRNAs 3 and 4 are conserved among the U.S. isolates (Fig. 1).

Analysis of the deduced amino acid sequences encoded by ORFs 2 to 5:

Fig. 2 shows the alignment of the deduced amino acid sequences of ORF 2 (A), ORF 3 (B), ORF 4 (C) and ORF 5 (D) of U.S. isolates ISU 22, ISU 55 and ISU 3927 with other known PRRSV isolates. The sequence of VR 2385 is shown on top, and only differences are indicated. Deletions are indicated by (-). The proposed signal peptide sequence in the ORF 5 of LV (D) is underlined (Meulenberg et al., 1995). Three hypervariable regions with antigenic potentials in ORF 5 (D) were indicated by asterisks (*). The published sequences used in this alignment were LV (Meulenberg et al., 1993), VR 2385 (U.S. Application Serial Nos. 08/131,625 and 08/301,435), VR 2332, ISU 79 and ISU 1894 (U.S. Application Serial No. 08/301,435).

On the basis of its high content of basic amino acids and its hydrophilic nature, the translation product of ORF 7 is

predicted to be the nucleocapsid protein (U.S. Application Serial Nos. 08/131,625 and 08/301,435; Meulenberg et al., 1993; Conzelmann et al., 1993; Mardassi et al., 1994). The ORF 6 product lacks a potential amino-terminal signal sequence and contains several hydrophobic regions which may represent the potential transmembrane fragments. Therefore, the ORF 6 product was predicted to be the M protein (U.S. Application Serial Nos. 08/131,625 and 08/301,435; Meulenberg et al., 1993; Conzelmann et al., 1993).

Computer analysis shows that the products encoded by ORFs 2 to 5 of the U.S. isolates all have hydropathy characteristics reminiscent of membrane-associated proteins. The translation products of ORFs 2 to 5 each contain a hydrophobic amino terminus. The N-terminal hydrophobic sequences may function as a signal sequence for each of these ORFs, and they may be involved in the transportation of ORFs 2 to 5 to the endoplasmic reticulum of infected cells. At least one additional hydrophobic domain in each of ORFs 2 to 5 was found at the carboxy termini. These additional hydrophobic domains may function as membrane anchors.

The deduced amino acid sequences of ORFs 2 to 5 of the seven U.S. isolates examined also varied considerably (Fig. 2), indicating that most of the nucleotide differences observed in Fig. 1 are not silent mutations. The amino acid sequence identity between VR 2385, VR 2332, ISU 22, ISU 55,

ISU 79, and ISU 1894 was 95-99% in ORF 2, 90-98% in ORF 3, 94-98% in ORF 4, and 88-97% in ORF 5 (Table 3).

Again, the least virulent isolate ISU 3927 displayed more variations with other U.S. isolates in ORFs 2 to 4 (Fig. 2 and Table 3) than in ORFs 5 to 7 (U.S. Application Serial No. 08/301,435 and Table 3). ORFs 2 to 5 of LV share only 57-61%, 55-56%, 65-67%, and 51-55% amino acid sequence identity with those ORFs of the U.S. isolates, respectively (Table 3). Deletions or insertions were found throughout ORFs 2 to 5 in comparing European LV and U.S. isolates (Fig. 2).

Sequence comparison of the ORF 5 product showed that the N-terminal region of ORF 5 is extremely variable, both (a) between U.S. isolates and LV and also (b) among the various U.S. isolates (Fig. 2D). In LV, the first 32-33 amino acid residues of ORF 5 may represent the signal sequence (Meulenberg et al., 1995; Fig. 2D). Therefore, the potential signal sequence of ORF 5 in all the PRRSV isolates is very heterogeneous. This heterogeneity is not due to any host immune selection pressure, because the signal peptide will be cleaved out and not be present in mature virions.

Three additional hypervariable regions were also identified by comparing the amino acid sequences of ORF 5 of all the PRRSV isolates available (Fig. 2D). Amino acid variations in these three regions are significant, and are not structurally conserved (Fig. 2D). Computer analysis indicates that all three hypervariable regions are hydrophilic and

antigenic. Thus, it is likely that these regions are exposed to the viral membrane and are under host immune selection pressure. However, further experiments may be necessary to confirm the specific functions of these hypervariable regions as antigenic determinants in the ORF 5 envelope protein.

The Phylogenetic relationships among U.S. isolates of PRRSV:

It has been shown previously that U.S. PRRSV and European PRRSV represent two distinct genotypes, based on analysis of the M and N genes (U.S. Application Serial No. 08/301,435). To determine the phylogenetic relationships of U.S. PRRSV isolates, ORFs 2 to 7 of the seven U.S. PRRSV isolates shown in Figs. 1 and 2 were first aligned with the GeneWorks program (intelligenetics, Inc.). The PAUP program (David L. Swofford, Illinois Natural History Survey, Champaign, IL) was then used to construct phylogenetic tree illustrating relationship among U.S. isolates of PRRSV.

The phylogenetic tree of Fig. 3 was constructed by maximum parsimony methods with the aid of the PAUP software package version 3.1.1. The branch with the shortest length (most parsimonious) was found by implementing the exhaustive search option. The branch lengths (numbers of amino acid substitutions) are given above each branch. The sequences used in the analysis are LV, VR 2385, VR 2332, ISU 79 and ISU 1894.

The phylogenetic tree indicates that at least three groups of variants (or minor genotypes) exist within the major U.S. PRRSV genotype. The least virulent U.S. PRRSV isolate ISU 3927 forms a branch distinct from other U.S. isolates (Fig. 3). Isolates ISU 22, ISU 79, ISU 1894, and VR 2332 form another branch, representing a second minor genotype. The third minor genotype is represented by isolates ISU 79 and VR 2385 (Fig. 3). A very similar tree was also obtained by analyzing the last 60 nucleotides of ORF 1b of the seven U.S. isolates presented in Fig. 1 (data not shown). Identical tree topology was also produced by the unweighted pair-group method with arithmetic mean (UPGMA) using the GeneWorks program (data not shown).

In summary, the different genotypes of PRRSV have been confirmed and further elucidated. At least three minor genotypes within the major genotype of U.S. PRRSV have been identified, based on an analysis of the sequence of ORFs 2 to 7. Genetic variations not only between the European PRRSV and the U.S. PRRSV but among the U.S. PRRSV isolates have also been further confirmed as well, indicating the heterogeneous nature of PRRSV. The least virulent U.S. PRRSV isolate ISU 3927 has unexpectedly high sequence variations in ORFs 2 to 4, as compared to other U.S. isolates.

Table 3: Nucleotide and deduced amino acid sequence identities (%) of ORFs 2 to 5 of PRRSV

ORF 2	VR2385	ISU22	ISU55	ISU79	ISU1894	ISU3927	VR2332	LV
VR2385	**	97	96	96	95	91	98	58
ISU22	97	**	96	98	96	93	99	59
ISU55	98	97	**	96	95	91	97	61
ISU79	96	97	97	**	96	91	98	60
ISU1894	96	97	96	96	**	93	96	57
ISU3927	94	94	94	93	93	**	93	58
VR2332	97	98	97	98	97	94	**	59
LV	65	66	66	67	66	65	66	**
ORF 3								
VR2385	**	91	94	92	90	87	91	55
ISU22	92	**	93	96	96	88	98	56
ISU55	94	93	**	94	93	87	94	56
ISU79	94	96	94	**	95	87	96	56
ISU1894	92	97	93	96	**	86	96	55
ISU3927	90	90	89	90	90	**	87	55
VR2332	93	98	94	97	97	90	**	56
LV	64	63	62	63	63	61	63	**
ORF 4								
VR2385	**	94	96	94	95	83	94	66
ISU22	93	**	94	97	99	93	98	66
ISU55	96	94	**	96	96	93	95	67
ISU79	93	97	94	**	98	92	96	66
ISU1894	92	98	94	96	**	93	98	66
ISU3927	91	93	92	91	91	**	92	67
VR2332	94	99	95	97	98	92	**	65
LV	66	66	63	65	66	65	65	**
ORF 5								
VR2385	**	90	91	88	89	91	89	54
ISU22	93	**	90	94	96	92	97	52
ISU55	94	92	**	89	89	90	89	51
ISU79	91	95	91	**	95	89	94	53
ISU1894	92	97	90	94	**	91	96	53
ISU3927	91	93	91	91	91	**	91	55
VR2332	93	98	91	95	97	92	**	53
LV	63	63	63	61	62	63	63	**

Note: The amino acid sequence comparisons are presented in the upper right half, and the nucleotide sequence comparisons are presented in the lower left half.

EXPERIMENT 2

During the replication of PRRSV, six subgenomic mRNAs (sg mRNAs), in addition to the genomic RNA, are synthesized. These sg mRNAs were characterized in this experiment.

The sg mRNAs of PRRSV form a 3'-coterminal nested set in PRRSV-infected cells. Each of these sg mRNAs is polycistronic and contains multiple open reading frames, except for sg mRNA 7 (as shown by Northern blot analysis using ORF-specific probes). The sg mRNAs were not packaged into virions, and only the genomic RNA was detected in purified virions, suggesting that the encapsidation signal of PRRSV is likely localized in the ORF 1 region.

The numbers of sg mRNAs in PRRSV-infected cells varies among PRRSV isolates with differing virulence. An additional species of sg mRNA in some PRRSV isolates was shown in Experiment 1 above to be derived from the sequence upstream of ORF 4, and has been designated as sg mRNA 4-1.

The leader-mRNA junction sequences of sg mRNAs 3 and 4 of isolates ISU 79 and ISU 1894, as well as sg mRNA 4-1 of the isolate ISU 79, contain a common six nucleotide sequence motif, T(G)TA(G/C)ACC. Sequence analysis of the genomic RNA of these two U.S. isolates and comparison with Lelystad virus (LV) revealed heterogeneity of the leader-mRNA junction

sequences among PRRSV isolates. The numbers, locations and the sequences of the leader-mRNA junction regions varied between U.S. isolates and LV, as well as among U.S. isolates. The last three nucleotides, ACC, of the leader-mRNA junction sequences are invariable. Variations were found in the first three nucleotides.

By comparing the 5'-terminal sequence of sg mRNA 4-1 with the genomic sequence of ISU 79 and ISU 1894, it was found that a single nucleotide substitution, from T in ISU 1894 to C in ISU 79, led to a new leader-mRNA junction sequence in ISU 79, and therefore, an additional species of sg mRNA (sg mRNA 4-1). A small ORF, designated as ORF 4-1, with a coding capacity of 45 amino acids was identified at the 5'-end of sg mRNA 4-1.

MATERIALS AND METHODS

Viruses and cells. The PRRSV isolates used (ISU 22, ISU 55, ISU 79, ISU 1894 and ISU 3927) were isolated from pig lungs obtained from different farms in Iowa. A continuous cell line, ATCC CRL 11171, was used for isolation and growth (culturing) of viruses. These PRRSV isolates were biologically cloned by three rounds of plaque purification and grown on the CRL 11171 cells. All of the virus isolates used in this study were at the seventh passage.

ISU 22 and ISU 79 are highly pathogenic and produce from 50 to 80% consolidation of the lung tissues in experimentally-infected five-week-old caesarean-derived colostrum-deprived

pigs necropsied at 10 days post-inoculation. By contrast, ISU 55, ISU 1894 and ISU 3927 are of low pathogenicity and produce only 10 to 25% consolidation of lung tissues in the same experiment (U.S. Application Serial Nos. 08/131,625 and 08/301,435).

Preparation of virus-specific total intracellular RNAs, poly (A)⁺ RNA and virion RNA. Confluent monolayers of CRL 11171 cells were infected with different isolates of PRRSV at the seventh passage at a multiplicity of infection (m.o.i.) of 0.1. PRRSV-specific total intracellular RNAs were isolated from PRRSV-infected cells by a conventional guanidinium isothiocyanate method (Stratagene). The poly (A)⁺ RNA was enriched from the total intracellular RNAs by oligo (dT)-cellulose column chromatography (Invitrogen).

For isolation of PRRSV virion RNA, confluent CRL 11171 cells were infected with isolate ISU 3927 of PRRSV at a m.o.i. of 0.1. When more than 70% of the infected cells showed a cytopathic effect, the cultures were frozen and thawed three times, and the culture medium was clarified at 1200 x g for 20 min. at 4 °C. The virus was then precipitated with polyethylene glycol and subsequently purified by cesium chloride gradient centrifugation as described in U.S. Application Serial No. 08/131,625. The purified virus was treated with RNase A at a final concentration of 20 μ/ml for 90 min. at 37°C. The virus was then pelleted, and the virion

RNA was isolated using a conventional guanidinium isothiocyanate method.

cDNA synthesis and polymerase chain reaction. cDNA was synthesized from total intracellular RNAs by reverse transcription using random primers and amplified by the polymerase chain reaction (RT-PCR) as described previously (Meng et al., 1993, J. Vet. Diagn. Invest., 5:254-258).

Northern blot analyses. Ten μg of total intracellular RNAs from virus infected cells and mock-infected cells were used per lane in a formaldehyde-agarose gel. For separation of poly (A)$^+$ RNA and virion RNA, fifteen ng of virion RNA and 0.2 μg of poly (A)$^+$ RNA were loaded per lane. The RNA was denatured with formaldehyde according to a conventional method (Sambrook et al, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). Electrophoretic separation of RNA, RNA blotting, and hybridization were performed as described in U.S. Application Serial No. 08/131,625. In some experiments, glyoxal-DMSO agarose gels were also performed as described in U.S. Application Serial No. 08/131,625.

For preparation of probes, a specific cDNA fragment from each of the ORFs 1b to 7 was generated by RT-PCR with ORF-specific primers. The primers were designed in such a way that each primer pair amplifies only a specific fragment of a given ORF, and the overlapping, neighboring ORFs are not included in any given cDNA probe. The primer pairs for

generating cDNA probes representing ORFs 1b through 7 are IM729/IM782 for ORF 1b, IM312/IM313 for ORF 2, XM1022/IM258 for ORF 3, XM1024/XMI 023 for ORF 4, PP287/PP286 for ORF 5, PP289/XM780 for ORF 6, and PP285/PP284 for ORF 7 (Table 4).

Cloning, sequencing and nucleotide sequence analyses. Primers for RT-PCR were designed on the basis of PRRSV isolate VR 2385 sequences, which amplified the entire protein coding regions of ORFs 2 to 5 of PRRSV isolates ISU 79 and ISU 1894. Primers JM259 and JM260 were used for amplification of ORFs 4 and 5, and XM992 and XM993 for amplification of ORFs 2 and 3 (Table 4). Unique restriction sites (EcoRI and BamHI) at the termini of the PCR products were introduced, thus enabling a cassette approach to replacement of these ORFs.

The PCR products of ORFs 2-3 and ORFs 4-5 of ISU 79 and ISU 1984 were each digested with EcoRI and BamHI, then purified and cloned into vector pSK+ as described previously (Meng et al., 1993, J. Vet. Diagn. Invest., 5:254-258). Plasmids containing viral inserts were sequenced with a conventional automated DNA sequencer (Applied Biosystem, Inc.). At least three cDNA clones representing the entire sequence of ORFs 2 to 5 from each virus isolate were sequenced with universal and reverse primers, as well as other virus-specific sequencing primers (XM969, XM970, XM1006, XM078 and XM077; see Table 4).

To determine the leader-mRNA junction sequences of sg mRNAs 3, 4 and 4-1, primer pair IM755 and DP586 (Table 4) was

used for RT-PCR to amplify the corresponding 5'-terminal sequences. The resulting PCR products were purified and sequenced by direct PCR sequencing using virus specific primers XMD77 and XM141 (Table 4). The sequences were combined and analyzed by MacVector (International Biotechnologies, Inc.) and GeneWorks (IntelliGenetics, Inc) computer software programs.

Oligonucleotides. The synthetic oligonucleotides used in this study were summarized in Table 4. These oligonucleotides were synthesized as single stranded DNA using an automated DNA synthesizer (Applied Biosystem) and purified by high pressure liquid chromatography (HPLC).

RESULTS

Sg mRNAs are not packaged into PRRSV virions. To determine whether the sg mRNAs of PRRSV are packaged, virions of PRRSV isolate ISU 3927 were purified by CsCl gradient. The purified virions were treated with RNase A before pelleting the virion and extracting RNA, to remove any RNA species which may have adhered to the virion surface. RNAs from RNase A-treated virions along with the total intracellular RNAs from isolate ISU 3927 of PRRSV-infected cells were separated in a formaldehyde gel and hybridized with a probe generated from the 3'-terminal sequence of the viral genome by PCR with primers PP284 and PP285 (U.S. Application Serial No. 08/131,625; Table 4).

Only the genomic RNA was detected in the purified virions of PRRSV isolate ISU 3927 (Fig. 4), and no detectable amounts of sg mRNAs were observed in the purified virions even after 3 weeks exposure. In contrast, seven species of sg mRNAs, in addition to the genomic RNA, were detected in ISU 3927-infected cells (Fig. 4). Similar results were observed with two other U.S. isolates, ISU 55 and ISU 79.

Variation in the numbers of the sg mRNAs among U.S. PRRSV isolates with differing virulence. All arteriviruses known prior to the present invention, including U.S. PRRSV and European PRRSV, have been shown to produce six sg mRNAs, except for three LDV variants (LDV-P, LDV-a and LDV-v), which synthesize seven sg mRNAs. However, a nested set of six sg mRNAs is produced in the LDV-C strain.

To compare if there are any variations in the sg mRNAs among U.S. PRRSV isolates, confluent monolayers of CRL 11171 cells were infected with five different isolates of U.S. PRRSV with differing virulence at a m.o.i. of 0.1. Total intracellular RNAs were isolated from virus-infected cells at 24 h post-infection. A cDNA fragment was generated from the extreme 3'-end of the viral genome by PCR with primers PP284 and PP285 (Table 4). The cDNA fragment was labelled with ^{32}P-dCTP by the random primer extension method, and hybridized with the total intracellular RNAs (separated on a formaldehyde gel).

Analyses of the RNAs showed that a nested set of six or more sg mRNAs, in addition to the genomic RNA, was present in cells infected with one of the five isolates of U.S. PRRSV with differing virulence (Fig. 5). Similar results were obtained when the total intracellular RNAs were separated on a glyoxal-DMSO agarose gel. PRRSV isolates ISU 55, ISU 79 and ISU 3927 produced seven easily distinguishable sg mRNAs, whereas isolates ISU 22 and ISU 1894 produced six sg mRNAs (Fig. 5). The U.S. PRRSV isolate VR 2385 also produces six sg mRNAs (U.S. Application Serial No. 08/131,625). An additional species of sg mRNA was located between sg mRNAs 3 and 4, and was designated as sg mRNA 4-1. The sg mRNAs differed little, if any, in size among the five isolates of PRRSV (Fig. 5). There appears to be no correlation, however, between the pneumovirulence and the numbers of the sg mRNAs observed in these five isolates.

Sg mRNA 4-1 is not a defective-interfering RNA and is not a result of nonspecific binding of the probes to ribosomal RNAs. It has been shown that, in coronaviruses, a variety of defective interfering RNA (DI RNA) of different sizes were generated when MHV was serially passaged in tissue culture at a high m.o.i. DI RNAs were also observed in cells infected with torovirus during undiluted passage. Therefore, the possibility of sg mRNA 4-1 of PRRSV being a DI RNA was investigated.

To exclude this possibility, the original virus stock of PRRSV isolate ISU 79, which produces the additional species of sg mRNA 4-1, was passaged four times in CRL 11171 cells at different m.o.i. of 0.1, 0.01 and 0.001, respectively. In a control experiment, four undiluted passages of the original virus stock of ISU 79 were performed. After four passages, total intracellular RNAs were isolated from virus-infected cells and Northern blot analysis was repeated with the same probe generated from the extreme 3'-end of the viral genome.

Analyses of the sg mRNAs showed that the additional species of sg mRNA 4-1 was still present in all RNA preparations with different m.o.i., as well as in RNA preparations from undiluted passages (Fig. 6A). Moreover, there was no interference or reduction in the synthesis of other sg mRNAs in the presence of sg mRNA 4-1, as is usually the case with DI RNA.

It has been demonstrated that the DI RNAs of MHV disappeared after two high-dilution passages. Therefore, if the original virus stock of ISU 79 contained DI RNA, then the DI RNA should disappear after four high-dilution passages. The experimental data above suggests that, unlike DI RNA, the replication of sg mRNA 4-1 is independent of the amount of standard virus. Thus, sg mRNA 4-1 is not a DI RNA.

In Northern blot analysis of total intracellular RNAS, the probes may nonspecifically bind to the 18S and 28S ribosomal RNAs, which are abundant in total cytoplasmic RNA

preparations. Alternatively, the abundant ribosomal RNAs may cause retardation of virus-specific sg mRNAs which may co-migrate corrugate with the ribosomal RNAs in the gel.

Two additional bands due to the nonspecific binding of probes to the ribosomal RNAs have been observed in LV-infected cells and LDV-infected cells. Therefore, it is possible that sg mRNA 4-1 of PRRSV is due to the nonspecific binding of probes to the ribosomal RNAs.

To rule out this possibility, polyadenylated RNA was isolated from total intracellular RNAs of CRL 11171 cells infected with either of two PRRSV isolates, ISU 55 and ISU 79. Both ISU 55 and ISU 79 produce the additional species of sg mRNA 4-1 (Fig. 5). Northern blot analysis of the polyadenylated RNA showed that the additional species of sg mRNA 4-1 in cells infected with either of these two isolates was still present (Fig. 6B), indicating that sg mRNA 4-1 is not due to the nonspecific binding of a probe to the ribosomal RNAS.

The sg mRNAs represent a 3'-coterminal nested set and the sg mRNA 4-1 is derived from the sequence upstream of ORF 4. Six sg mRNAs, in addition to the genomic RNA, are detected in cells infected with VR 2385 using a cDNA probe from the extreme 3'-end of the viral genome (U.S. Application Serial No. 08/131,625). Thus, like Berne virus (BEV), LDV, EAV, coronaviruses and LV, the replication of U.S. PRRSV also requires the synthesis of a 3'-coterminal nested set of sg

mRNAs (U.S. Application Serial Nos. 08/131,625 and 08/301,435).

To analyze these sg mRNAs in more detail, seven cDNA fragments specific for each of ORFs 1b through 7 were amplified by PCR. The design of primers for PCR was based on the sequence of VR 2385. The sequences and locations of the primers, IM729 and IM782 for ORF 1b, IM312 and IM313 for ORF 2, XM1022 and IM258 for ORF 3, XM1024 and XM1023 for ORF 4, PP286 and PP287 for ORF 5, PP289 and XM780 for ORF 6, and PP284 and PP285 for ORF 7 and the 3′ noncoding region (NCR), are shown in Table 4. The primers were designed in such a way that each set of primers will only amplify a fragment from a particular ORF, and the overlapping sequences between neighboring ORFs are not included in any given fragment. Therefore, each of these seven DNA fragments represents only one particular ORF except for fragment 7, which represents both ORF 7 and the 3′-NCR.

These seven DNA fragments were labeled with ^{32}P-dCTP and hybridized to Northern blots of total intracellular RNAs extracted from cells infected with either of two U.S. isolates of PRRSV, ISU 1894 and ISU 79. Total intracellular RNAs isolated from mock-infected CRL 11171 cells were included as a control.

Northern blot analyses showed that Probe 1, generated from ORF 1b, hybridized only with the genomic RNA. Probes 2 through 7 each hybridized with one more additional RNA species

besides the genomic RNA (Fig. 7). The results indicate that a 3'-coterminal nested set of six (ISU 1894) or more (ISU 79) sg mRNAs is formed in PRRSV-infected cells (Figs. 7A and 7B), with the smallest 3'-terminal RNA (sg mRNA 7) encoding ORF 7. The sg mRNAs of U.S. PRRSV all contain the 3'-end of the genomic RNA, but extend for various distances towards the 5'-end of the genome, depending on the size of the given sg mRNA.

The sg mRNA 4-1 of PRRSV isolate ISU 79 hybridized with probes 4 through 7, but not with probes 1, 2 and 3 (Fig. 7B), suggesting that sg mRNA 4-1 contains ORFs 4 through 7 as well as the 3'-NCR. Therefore, sg mRNA 4-1 is generated from the sequence upstream of ORF 4.

A single nucleotide substitution leads to the acquisition of the additional species of sg mRNA 4-1. Northern blot hybridization data showed that sg mRNA 4-1 is derived from the sequence upstream of ORF 4 (Fig. 7B). To determine the exact location and the leader-mRNA junction sequence of sg mRNA 4-1, a set of primers, IM755 and DP586, was designed (Table 4). The forward primer IM755 was based on the 3'-end of the leader sequence of VR 2385, and the reverse primer DP586 is located in ORF 4 (Table 4).

RT-PCR with primers IM755 and DP586 was performed using total intracellular RNAs isolated from cells infected with either of ISU 1894 or ISU 79. ISU 79 produces sg mRNA 4-1, but ISU 1894 does not (Fig. 5). A 30-second PCR extension time was applied to preferentially amplify the short fragments

representing the 5'-terminal sequences of sg mRNAs 3, 4 and 4-1.

Analysis of the RT-PCR products showed that two fragments with sizes of about 1.1 kb and 0.45 kb were amplified from the total RNAs of ISU 1894 virus-infected cells (Fig. 8A). These two fragments represent 5'-portions of sg mRNAs 3 and 4, respectively. In addition to the two fragments observed in the isolate of ISU 1894, a third fragment of about 0.6 kb representing the 5'-portion of sg mRNA 4-1 was also amplified from total RNAs of cells infected with ISU 79 (Fig. 8A).

To determine the leader-mRNA junction sequences of sg mRNAs 3, 4 and 4-1, the RT-PCR products of ISU 79 and ISU 1894 were purified from an agarose gel using a GENECLEAN kit (Bio 101, Inc.), and sequenced directly with an automated DNA Sequencer (Applied Biosystems). The primers used for sequencing the 5'-end of the RT-PCR products (XM141 and XM077, Table 4) were designed on the basis of the genomic sequences of ISU 79 and ISU 1894 (Fig. 9). The leader-mRNA junction sequences (in which the leader joins the mRNA body during the synthesis of sg mRNAs) of sg mRNAs 3, 4, and 4-1 of the two U.S. PRRSV isolates were determined by comparing the sequences of the 5'-end of the sg mRNAs and the genomic RNA of the two isolates (Fig. 8B).

The leader-mRNA junction sequences of sg mRNAs 3 and 4 of ISU 1894 and ISU 79 were identical. For sg mRNA 3, the leader-junction sequence (GUAACC) is located 89 nucleotides

upstream of ORF 3. For sg mRNA 4, UUCACC is located 10 nucleotides upstream of ORF 4 (Fig. 8B and Fig. 9). The leader-mRNA junction sequence of sg mRNA 4-1 of ISU 79 is UUGACC, located 236 nucleotides upstream of ORF 4 (Figs. 8B and 9).

Sequence alignment of the genomic sequences of ISU 79 and ISU 1894 shows that a single nucleotide substitution, from T in ISU 1894 to C in ISU 79, leads to the acquisition of an additional leader-mRNA junction sequence, UUGACC, in ISU 79 (Figs. 8B and 9). Therefore, an additional species of sg mRNA (4-1) is formed (Fig. 5). In addition to ORFs 4 to 7 contained within sg mRNA 4, sg mRNA 4-1 contains at the 5'-end an additional small ORF (ORF 4-1) with a coding capacity of 45 amino acids (Fig. 9). This small ORF stops just one nucleotide before the start codon of ORF 4.

Sequence analyses of ORFs 2 to 7 of two U.S. isolates reveal heterogeneity of the leader mRNA junction sequences. ORFs 2 to 5 of ISU 79 and ISU 1894 were cloned and sequenced (see Experiment 1 above). ISU 79 produces seven easily distinguishable sg mRNAs, whereas ISU 1894 produces six distinguishable sg mRNAs (Figs. 5 and 7). At least three cDNA clones at any given region of ORFs 2 to 5 were sequenced for each virus isolate, using universal and reverse primers as well as virus-specific primers XM969, XM970, XM1006, XM078, and XM077 (Table 4). The sequences of ORFs 6 and 7 of ISU

1894 and ISU 79 are disclosed in U.S. Application Serial No. 08/301,435.

Sequence analysis showed that the ORFs 2 to 7 of ISU 79 and ISU 1894 overlap each other except for a 10-nucleotide noncoding region between ORF 4 and ORF 5. The same observation was previously made for VR 2385 (U.S. Application Serial No. 08/301,435). This is very unusual, since all members of the proposed *Arteriviridae* family, including LV, contain overlapping ORFs. However, the ORFs of coronaviruses are separated by intergenic noncoding sequences. Therefore, U.S. PRRSV appears to be somewhat similar to the coronaviruses in terms of the genomic organization in junction regions of ORFs 4 and 5.

ORF 2 of ISU 1894 was one amino acid longer than that of ISU 79 (Fig. 9). The stop codon of ORF 2, TAG, was changed to TGG in ISU 1894 immediately followed by a new stop codon (TGA) in ISU 1894 (Fig. 9). The sizes of other ORFs of ISU 79 and ISU 1894 were identical (Fig. 9). There were no deletions or insertions in ORFs 2 to 7 of these isolates. However, numerous substitutions are present throughout the entire sequence of ORFs 2 to 7 between ISU 79 and ISU 1894 (Fig. 9).

The numbers and locations of the determined or predicted leader-mRNA junction sequences varied between ISU 1894 and ISU 79 (Fig. 9). In addition to the regular leader-mRNA 4 junction sequence, TTCACC, 10 nucleotides upstream of ORF 4, there was an additional leader-mRNA 4-1 junction sequence

(TTGACC) located 236 nucleotides upstream of ORF 4 in ISU 79 (Fig. 9). The leader-mRNA junction sequences of sg mRNAs 4 and 4-1 were separated by 226 nucleotides, which correlated with the estimated sizes of sg mRNAs 4 and 4-1 observed in Northern blot analysis (Fig. 5) and RT-PCR amplification (Fig. 8A).

The leader-mRNA 3 junction sequence is identical between ISU 1894 and ISU 79, GTAACC, located 89 nucleotides upstream of ORF 3. The predicted leader-mRNA junction sequences of sg mRNAs 2 and 6 of ISU 1894 and ISU 79 were also the same (Fig. 9).

However, the predicted leader-mRNA 5 junction sequences of ISU 1894 and ISU 79 are different (Fig. 9). There are 3 potential leader-mRNA 5 junction sequences for ISU 79 (GCAACC, GAGACC and TCGACC, located 55, 70 and 105 nucleotides upstream of ORF 5, respectively). Two potential leader-mRNA 5 junction sequences were also found in ISU 1894 (GAAACC and TCGACC, located 70 and 105 nucleotides upstream of ORF 5, respectively) (Fig. 9). The differences were due to the two-nucleotide substitutions in the predicted leader-mRNA 5 junction sequences of these isolates (Fig. 9).

In addition to the leader-mRNA 7 junction sequence 15 nucleotides upstream of ORF 7, an additional leader-mRNA 7 junction sequence was found (ATAACC), located 129 nucleotides upstream of ORF 7 in each of these two isolates (Fig. 9). However, the sg mRNA corresponding to this additional leader-

mRNA 7 junction sequence was not clearly distinguishable from the abundant sg mRNA 7 which produced a widely-diffused band in the Northern blot (Figs. 5, 6 and 7).

Variations in the numbers and locations of the leader-mRNA junction sequences between LV and the two U.S. isolates analyzed in this experiment were also found by comparing the leader-mRNA junction sequences of LV with those of the two U. S. isolates ISU 1894 and ISU 79. Taken together, these data indicate that the sg mRNAs of PRRSV are polymorphic, and the numbers and the exact sizes of the sg mRNAs depend on the particular PRRSV isolate analyzed. However, a nested set of six sg mRNAs most likely reflects the standard arterivirus genome organization and transcription.

Table 4. Synthetic oligonucleotides used in Experiment 2

Oligo Name	Sequence	Location (nucleotides)[a]	Polarity[b]
IM729	5'-GACTGATGGTCTGGAAAG-3'	ORF1b, -507 to -490 upstream of ORF2	+
IM782	5'-CTGTATCCGATTCAAACC-3'	ORF1b, -180 to -163 upstream of ORF2	-
IM312	5'-AGGTTGGCTGGTGGTCTT-3'	ORF2, 131 to 148 downstream of ORF2	+
IM313	5'-TCGCTCACTACCTGTTTC-3'	ORF2, 381 to 398 downstream of ORF2	-
XM1022	5'-TGTGCCCGCCTTGCCTCA-3'	ORF3, 168 to 175 downstream of OEF3	+
IM268	5'-AAACCAATTGCCCCCGTC-3'	ORF3, 520 to 537 downstream of ORF3	-
XM1024	5'-TATATCACTGTCACAGCC-3'	ORF4, 232 to 249 downstream of ORF4	+
XM1023	5'-CAAATTGCCAACAGAATG-3'	ORF4, 519 to 536 downstream of ORF4	-
PP287	5'-CAACTTGACGCTATGTGAGC-3'	ORF5, 129 to 148 downstream of ORF5	+
PP286	5'-GCCGCGGAACCATCAAGCAC-3'	ORF5, 538 to 667 downstream of ORF5	-
PP289	5'-GACTGCTAGGGCTTCTGCAC-3'	ORF6, 119 to 138 downstream of ORF6	+
XM780	5'-CGTTGACCGTAGTGGAGC-3'	ORF6, 416 to 433 downstream of ORF6	-
PP285	5'-CCCCATTTCCCTCTAGCGACTG-3'	ORF7, 157 to 178 downstream of ORF7	+
PP284	5'-CGGCCGTGTGGTTCTCGCCAAT-3'	3' NCR, -27 to -6 upstream of poly (A)	-
JM260	5'-GGGGAATTCGGGATAGGGAATGTG-3'	ORF3, 338 to 356 downstream of ORF3	+
JM259	5'-GGGGATCCTTTTGTGGAGCCGT-3'	ORF6, 34 to 52 downstream of ORF6	-
XM993	5'-GGTGAATTCGTTTTATTTCCCTCCGGGC-3'	ORF1b, -53 to -35 upstream of ORF2	+
XM992	5'-GGGGGATCCTGTTGGTAATAG/AGTCTG-3'	ORF3, -50 to -34 upstream of ORF4	-
XM970	5'-GGTTTCACCTAGAATGGC-3'	ORF2, 522 to 550 downstream of ORF2	+
XM969	5'-GATAGAGTCTGCCCTTAG-3'	ORF5, 443 to 460 downstream of ORF6	-
XM1006	5'-GCTTCTGAGATGAGTGA-3'	ORF4, 316 to 332 downstream of ORF4	+
XM078	5'-CTGAGCAATTACAGAAG-3'	ORF2, 202 to 218 downstream of ORF2	+
XM077	5'-CAACCAGGCGTAAACACT-3'	ORF3, 316 to 333 downstream of ORF3	-
IM755	5'-GACTGCTTTACGGTCTCTC-3'	Leader, 3' end of the leader sequence	+
DP586	5'-GATGCCTGACACATTGCC-3'	ORF4, 355 to 372 downstream of ORF4	-
XM141	5'-CTGCAAGACTCGAACTGAA-3'	ORF4, 78 to 97 downstream of ORF4	-

a. The oligonucleotides were designed on the basis of sequence data presented in this application and U.S. Application Serial Nos. 08/131,625 and 08/301,435

b. Oligonucleotides complementary to the genomic RNA have negative (-) polarities.

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.

WHAT IS CLAIMED AS NEW AND DESIRED TO BE SECURED BY LETTERS PATENT OF THE UNITED STATES IS:

1. A purified preparation comprising a polynucleic acid encoding at least one polypeptide selected from the group consisting of:

proteins encoded by one or more open reading frames (ORF's) of an Iowa strain of porcine reproductive and respiratory syndrome virus (PRRSV);

proteins at least 99% but less than 100% homologous with those encoded by one or more of ORFs 2 of VR 2385 and VR 2474, ORF 3 of VR 2429, and ORFs 4 of VR 2429 and ISU 1894;

proteins at least 98% but less than 100% homologous with those encoded by one or more of ORF 2 of VR 2430 and ORF 5 of VR 2429;

proteins at least 97% but less than 100% homologous with those encoded by one or more of ORF 2 of ISU 1894, ORFs 3 of VR 2474 and ISU 1894, ORF 4 of VR 2474 and ORF 5 of ISU 1894;

proteins at least 96% but less than 100% homologous with the protein encoded by ORF 4 of VR 2430;

proteins at least 95% but less than 100% homologous with those encoded by one or more of ORF 3 of VR 2430, ORF 4 of VR 2385, and ORF 5 of VR 2474;

proteins at least 94% but less than 100% homologous with the protein encoded by ORF 2 of VR 2431;

proteins at least 93% but less than 100% homologous with the protein encoded by ORF 3 of VR 2431;

proteins at least 92% but less than 100% homologous with the proteins encoded by one or more of ORF 3 of VR 2385 and ORF 5 of VR 2431;

proteins at least 90% but less than 100% homologous with those encoded by one or more of ORFs 5 of VR 2385 and VR 2430;

proteins at least 88% but less than 100% homologous with the protein encoded by ORF 3 of VR 2431;

proteins at least 97% but less than 100% homologous with proteins encoded by one or both of ORF 6 and ORF 7 of an Iowa strain of PRRSV;

antigenic regions of said proteins which are at least 5 amino acids in length;

proteins encoded by one or more of ORFs 2, 3 and 4 of a PRRSV other than VR 2431 which include a low-virulence mutation, or a low-virulence fragment thereof at least 5 amino acids in length;

and combinations thereof.

2. The purified preparation of Claim 1, wherein said polynucleic acid has a sequence selected from the group consisting of the formulas (I), (II), (III) and (IV):

$$5'-\alpha-\beta-3' \qquad \text{(I)}$$

$$5'-\alpha-\beta-\gamma-3' \quad \text{(II)}$$

$$5'-\beta-\delta-\gamma-3' \quad \text{(III)}$$

$$5'-\alpha-\beta-\delta-\gamma-3' \quad \text{(IV)}$$

wherein:

α encodes at least one polypeptide, low-virulence mutant thereof, or antigenic or low-virulence fragment thereof encoded by a polynucleotide selected from the group consisting of ORFs 2, 3 and 4 of an Iowa strain of PRRSV;

β is at least one copy of an ORF 5 from an Iowa strain of PRRSV or an antigenic fragment thereof (e.g. one or more hypervariable regions);

γ encodes at least one polypeptide encoded by a polynucleotide selected from the group consisting of ORF 6 and ORF 7 of an Iowa strain of PRRSV, or an antigenic fragment thereof; and

δ is either a covalent bond or a linking polynucleic acid which does not materially affect transcription and/or translation of the polynucleic acid.

3. A purified preparation comprising a polynucleic acid having the formula (V):

$$5'-\epsilon-\zeta-\iota-\kappa-\xi-3' \quad \text{(V)}$$

where:

κ is a polynucleotide comprising one or more genes selected from the group consisting of (a) a conventional marker or reporter gene, (b) genes encoding at least one polypeptide, or antigenic or low-virulence fragment thereof, encoded by one or more of ORFs 2, 3 and 4 of an Iowa strain of PRRSV; (c) genes encoding at least one polypeptide, or antigenic fragment thereof, encoded by an ORF 5 from an Iowa strain of PRRSV; (d) genes encoding at least one polypeptide, or an antigenic fragment thereof, encoded by an ORF 6 or ORF 7 of an Iowa strain of PRRSV; and (e) operationally linked combinations thereof;

ϵ, which is optionally present, is a 5'-terminal polynucleotide sequence which provides a means for operationally expressing the polynucleotide κ;

ζ is a polynucleotide of the formula KTVACC, where K is T, G or U, and V is A, G or C;

ι is a polynucleotide of at most about 130 nucleotides in length;

and ξ, which is optionally present and which may be operationally linked to ϵ, is a 3'-terminal polynucleotide sequence which does not suppress the operational expression of the polynucleotide κ.

4. The purified preparation of Claim 1, wherein said polynucleotide encodes at least one hypervariable region of a protein encoded by an ORF 5 of an Iowa strain of PRRSV.

5. The purified preparation of Claim 1, wherein said polypeptide is selected from the group consisting of proteins at least 97% homologous with those encoded by ORF 6 of VR 2385, VR 2429 (ISU-22), ISU-79 and VR 2431 (ISU-3927); proteins at least 90% homologous with proteins encoded by ORFs 5 of VR 2385 and VR 2430; proteins at least 94% homologous with the protein encoded by ORF 2, at least 88% homologous with the protein encoded by ORF 3, and at least 93% homologous with the protein encoded by ORF 4 of VR 2431; and antigenic regions of said proteins having a binding affinity to an antibody specific to the corresponding full-length protein of at least 1% of the binding affinity of the corresponding full-length protein.

6. The purified preparation of Claim 5, wherein isolated polynucleic acid is selected from the group consisting of ORFs 2, 3, 4, 5, 6 and 7 of VR 2385 and VR 2431.

7. The purified preparation of Claim 1, wherein said polynucleic acid encodes said homologous protein, and non-homologous residues in said homologous protein are conservatively substituted.

8. A purified polypeptide encoded by the polynucleic acid of Claim 1.

9. A purified polypeptide encoded by the polynucleic acid of Claim 5.

10. A vaccine, comprising (a) an effective amount of the polypeptide of Claim 8 to protect a pig against a porcine reproductive and respiratory syndrome, and (b) a physiologically acceptable carrier.

11. A vaccine, comprising (a) an effective amount of the polynucleic acid of Claim 1 to protect a pig against porcine reproductive and respiratory syndrome, and (b) a physiologically acceptable carrier.

12. A method of protecting a pig from porcine reproductive and respiratory syndrome, comprising administering an effective amount of the vaccine of Claim 11 to a pig in need thereof.

13. The method of Claim 12, wherein said vaccine is administered orally or parenterally.

14. The method of Claim 13, wherein said vaccine is administered intramuscularly, intradermally, intravenously, intraperitoneally, subcutaneously or intranasally.

15. The method of Claim 12, wherein said vaccine is administered to a sow in need thereof.

16. An antibody which specifically binds to the polypeptide of Claim 9.

17. The antibody of Claim 16, wherein said antibody is a monoclonal antibody.

18. A method of treating a pig suffering from porcine reproductive and respiratory syndrome, comprising administering an effective amount of the antibody of Claim 16 to a pig in need thereof.

19. A diagnostic kit for assaying a porcine reproductive and respiratory syndrome virus, comprising the antibody of Claim 16 and a diagnostic agent which indicates a positive immunological reaction with said antibody.

20. A method of producing a polypeptide, comprising expressing the polynucleic acid of Claim 1 in an operational expression system, and purifying said expressed polypeptide from said expression system.

21. A method of producing a polynucleic acid, comprising (i) digesting cDNA complementary to genomic

polynucleic acids of an Iowa strain of PRRSV with one or more restriction enzymes; (ii) amplifying genomic polynucleic acids of an Iowa strain of PRRSV by PCR (using appropriate primers complimentary to the 5′ and 3′-terminal regions of the desired ORF(s) or to regions upstream of the 5′-terminus or downstream from the 3′-terminus) and cloning, or (iii) synthesizing using a polynucleotide synthesizer.

```
                                    1/21   +1>ORF2
VR2385   CCTGTCATTGAACCAACTTTAGGCCTGAATTGAGATGAAATGGGGTCTATGCAAAGCCTT   60
ISU3927  ...A.............G..T..AG.C...A...C.........C.............   60
ISU55    ...A.................................A.....................   60
ISU22    .....................................A.....G.C.............   60
VR2332   .....................................A.....C.............   60
ISU1894  ..C..................................A.....CG............   60
ISU79    .....................................A.....C.............   60

VR2385   TTTGACAAAATTGGCCAACTTTTTGTGGATGCTTTCACGGAGTTCTTGGTGTCCATTGTT  120
ISU3927  ........G..C..T.....................................................  120
ISU55    ........................C.....................................  120
ISU22    .....................................................  120
VR2332   .....................................................  120
ISU1894  .....................................................  120
ISU79    .....................................................  120

VR2385   GATATCATTATATTTTTGGCCATTTTGTTTGGCTTCACCATCGCAGGTTGGCTGGTGGTC  180
ISU2927  ........C.....................................T..C..C.............  180
ISU55    .....................................C.....................  180
ISU22    .....................................C.....................  180
VR2332   .....................................C.....................  180
ISU1894  ........C.....................................C.....................  180
ISU79    ..............C.....................................C.....................  180

VR2385   TTTTGCATCAGATTGGTTTGCTCCGCGATACTCCGTGCGCGCCCTGCCATTCACTCTGAG  240
ISU3297  .....................................................C.....  240
ISU55    ............................C.....................................  240
ISU22    .....................................................  240
VR2332   .....................................A.....................  240
ISU1894  .....................................A.....................  240
ISU79    .....................................A.....................  240

VR2385   CAATTACAGAAGATCCTATGAGGCCTTTCTCTCTCAGTGCCAGGTGGACATTCCCACCTG  300
ISU3297  ..............................T.....................G....  300
ISU55    ..............................T.....................  300
ISU22    ..............T.............T..C.....A.....................  300
VR2332   ..............T.............T..C.....A.....................  300
ISU1894  ..............T...........T.A..C.....A.....................  300
ISU79    ..............T.............T..C.....A.....................  300

VR2385   GGGAACTAAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGA  360
ISU3927  ......A.G.........A.................C.....................  360
ISU55    .....T.....................T.....................  360
ISU22    ........................T.G.....................  360
VR2332   .....................................................  360
ISU1894  ........................T.....................  360
ISU79    ........................T.....................  360
```

FIG. 1A

```
VR2385   AATGGTGTCGCGTCGAATGTACCGCATCATGGAAAAAGCAGGACAGGCTGCCTGGAAACA  420
ISU3927  ............................................................  420
ISU55    ............................................................  420
ISU22    ...........................................G................  420
VR2332   ...........................................G................  420
ISU1894  ..............C............................G................  420
ISU79    G...........................................................  420

VR2385   GGTAGTGAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGGATGTGGTGGCTCATTTTCA  480
ISU3927  ...G.....T.........................G..............C........  480
ISU55    ...G........................................................C.....  480
ISU22    ...G........................................................  480
VR2332   ...G........................................................  480
ISU1894  ...G.....................C......................................  480
ISU79    ...G........................................................  480

VR2385   GCATCTTGCCGCCATTGAAGCCGAGACCTGTAAATATCTGGCCTCTCGGCTGCCCATGCT  540
ISU3927  ...C.....................T.........T..........T..........  540
ISU55    ...........................................T..............  540
ISU22    .........T.................................T.......C.....  540
VR2332   ......A....................................T.......C.....  540
ISU1894  ..........................T.........T.......C.....  540
ISU79    ..............C.....................T.......C.....  540
                                  -89(mRNA3)
VR2385   ACACCACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTATAATAGTACTTTGAATCA  600
ISU3927  ....A.......T..T.....______.....C.....  600
ISU55    ....A.....______.....  600
ISU22    ....A.....______.....  600
VR2332   ....A.....______.....C.....  600
ISU1894  ...TA.....A.....______.....C.G.G..  600
ISU79    ....A.....______.....G  600
                                                                  +1>
VR2385   GGTGTTTGCTGTTTTCCCAACCCCTGGTTCCCGGCCAAAGCTTCATGATTTCCAGCAATG  660
ISU3927  ..........A.....C.....T.....C.....T.....  660
ISU55    ....C.....A.....T.....  660
ISU22    ..........A.....T.....  660
VR2332   ..........A....T.....T.....  660
ISU1894  ..........A.....T.....  660
ISU79    ..........A.....C..T.....  660
         ORF3
VR2385   GCTAATAGCTGTACATTCCTCTATATTTTCCTCTGTTGCAGCTTCTTGTACTCTTTTTGT  720
ISU3927  .....C....CC.....  720
ISU55    ............................................................  720
ISU22    .T.....C.....C.....  720
VR2332   .T.....C.....  720
ISU1894  .T.....C.....  720
ISU79    ..........G.....C.....C.....  720
```

FIG. 1B

```
VR2385    TGTGCTGTGGTTGCGGGTTCCAATGCTACGTACTGTTTTTGGTTTCCGCTGGTTAGGGGC  780
ISU3927   ....T.......A...A.G.............T.............................  780
ISU55     ..................................T..C.......................  780
VR2385    ....................................................................  780
ISU22     ........................A.....T.....................................  780
VR2332    ........................A.....................................  780
ISU1894   ..........C.....................................................G  780
ISU79     ........................A.....................................  780
                                  **<ORF2
VR2385    AATTTTTCTTTCGAACTCACGGTGAATTACACGGTGTGCCCGCCTTGCCTCACCCGGCAA  840
ISU3927   ........C.....G....T.....................A.................G  840
ISU55     ....................A.....C...........T..A.....................  840
VR2385    ....................................................................  840
ISU22     ..............G.....................T..A.....................  840
VR2332    ....................A.................T..A.....................  840
ISU1894   ........C.........T.................T..A.....................  840
ISU79     ....................TA.....................A.....................  840
                                  ***
VR2385    GCAGCCGCAGAGGCCTACGAACCCGGCAGGTCCCTTTGGTGCAGGATAGGGCATGATCGA  900
ISU3927   .........C..AT..........AA......T..............C...A.........  900
ISU55     ......A.................T........T.....................T.......C  900
ISU22     ............AT............T.....T.....................T....C...  900
VR2332    ......A.....AT............T.....T.....................T....C...  900
ISU1894   ............T............T.....T.....................T....C...  900
ISU79     ........................T.....T.....................T.C.......  900

VR2385    TGTGGGGAGGACGATCATGATGAACTAGGGTTTGTGGTGCCGTCTGGCCTCTCCAGCGAA  960
ISU3927   .....T...........C..C........A...ACA..A...C.............AA...  960
ISU55     ....................C.....................................  960
ISU22     ....................C..G.........A....A..AC.....T..........  960
VR2332    ....................C..G.........A..A.A...C.................  960
ISU1894   ....................C..G.........A..A.A...C.................  960
ISU79     .....A.........C.....C..G.........A..A.A.....................  960
                  -236(mRNA4-1)
VR2385    GGCCACTTGACCAGTGCTTACGCCTGGTTGGCGTCCCTGTCCTTCAGCTATACGGCCCAG  1020
ISU3927   .T...T_____....T.................TT...........T..C.........  1020
ISU55     ......_____....T.................T.....T.....T..C..A......  1020
ISU22     ..........T....T.................TTT.............C.........  1020
VR2332    ..........TG...T.................T.T.............C.........  1020
ISU1894   ..........T....T.................TTT.............C.........  1020
ISU79     ......_____...T.................T.T.............C.........  1020
                                                    +1>ORF4-1
VR2385    TTCCATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTCTATGTTGACATCAAGCAC  1080
ISU3927   ..T.......................A.........AAG.....___.............T..T  1080
ISU55     ........T.....................................A...T.___.........G...T  1080
ISU22     ..............................C.............T.___.............A..T  1080
VR2332    ....................................................T.___.....A..T  1080
ISU1894   ....................................................T.___.....A..T  1080
ISU79     .....C..............................................T.___.....A..T  1080
```

FIG. 1C

```
VR2385   CAATTCATTTGCGCTGTTCATGATGGGCAGAACACCACCTTGCCCCACCATGACAACATT 1140
ISU3927  ...C.......T...........C.....................T.G.............. 1140
ISU55    ..............C.....C..C.........G....T.....T.G.........T.... 1140
ISU22    ...C....C.....C.AA.....C.....A..............T.GT.............. 1140
VR2332   ...C....C.....C.AA.....C.....................T.GT.............. 1140
ISU1894  ...C..........C.AA.....C..A......G..........T.GT........T.... 1140
ISU79    ...C....C.....C.AA.....C.....................T.GT.............. 1140
                                                          -10(mRNA4)
VR2385   TCAGCCGTGCTTCAGACCTATTACCAGCATCAGGTCGACGGGGGCAATTGGTTTCACCTA 1200
ISU3927  ..T......T................A..C..A.....T..T.....C.....______.... 1200
ISU55    .........T.C.....T........A.....A........C.....______.... 1200
ISU22    .........T................A.....A........T.....______.... 1200
VR2332   .........T................A.....A........C.....______.... 1200
ISU1894  .........T................A.....A.....T..C.....______.... 1200
ISU79    ..G......T................A.....A........C.....______.... 1200
           +1>ORF4                                                **
VR2385   GAATGGGTGCGTCCCTTCTTTTCCTCTTGGTTGGTTTTAAATGTCTCTTGGTTTCTCAGG 1260
ISU3927  ......C.................................G........C.............. 1260
ISU55    ......C.................................................... 1260
ISU22    ......C.T................A.................................... 1260
VR2332   ......C.T................G.................................... 1260
ISU1894  ......C.T................A.....................A.......... 1260
ISU79    ......C.................A.................................... 1260
         *<ORF4-1
VR2385   CGTTCGCCTGCAAGCCATGTTTCAGTTCGAGTCTTTCAGACATCAAGACCAACACCACCG 1320
ISU3927  ...................................................................... 1320
ISU55    ..................................G.......T.................. 1320
ISU22    ............A.....................G....T..T.................. 1320
VR2332   ............A.....................G....T..T.................. 1320
ISU1894  ............A.....C.............G....T..T..........T..... 1320
ISU79    ............A.....................G.......T.................. 1320

VR2385   CAGCGGCAGGCTTTGCTGTCCTCCAAGACATCAGTTGCCTTAGGCATCGCAACTCGGCCT 1380
ISU3927  .G..A...AAT..C..........G.....G.C.............G........A..A 1380
ISU55    ...........................................T.................. 1380
ISU22    ........A.....................G.................G.......... 1380
VR2332   ........A.....................................G.......... 1380
ISU1894  ........A.....................................................  1380
ISU79    ........A.....................................................  1380
                                                  ***<ORF3
VR2385   CTGAGGCGATTCGCAAAGTCCCTCAGTGCCGCACGGCGATAGGGACACCCGTGTATATCA 1440
ISU3927  ........T........A.....................................................  1440
ISU55    ........T........A..........T..T.....................A.......T.  1440
ISU22    ..................A.............T.........................T.  1440
VR2332   ..................A.............T.....................G.T.  1440
ISU1894  ..................A.............T.........................T.  1440
ISU79    ..................A.............T.................TA.....T.  1440
```

FIG. 1D

```
VR2385    CTGTCACAGCCAATGTTACCGATGAGAATTATTTGCATTCCTCTGATCTTCTCATGCTTT 1500
ISU3927   ..A.....A..A.....C.....T..... 1500
ISU55     ..G.....A.....C..... 1500
ISU22     .CA.....G..A.....A....T.....C..... 1500
VR2332    .CA.....G..A.....A....T.....C..... 1500
ISU1894   .CA.....G..A.....A....T.....C..... 1500
ISU79     .CA.....G..A....A.....A....T.....C.....C. 1500

VR2385    CTTCTTGCCTTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTTAAGGTGGTATTTGGCA 1560
ISU3927   .C.....C.....G.....G..... 1560
ISU55     .....C.....A..... 1560
ISU22     .....G..... 1560
VR2332    ..... 1560
ISU1894   .....C.....A....C..... 1560
ISU79     .....G.....T..... 1560

VR2385    ATGTGTCAGGCATCGTGGCAGTGTGCGTCAACTTCACCAGTTACGTCCAACATGTCAAGG 1620
ISU3927   .....C.....T.....A....T..T.....C..T.....C..... 1620
ISU55     .....T.....T.....C..... 1620
ISU22     .....T.....T.....T..T.....C..T.....G.. 1620
VR2332    .....T.....T.....T..T.....C..... 1620
ISU1894   .....T.....T.....T..T.....C.....G.. 1620
ISU79     .....T.....T.....T..T.....C.....T.....G.. 1620

VR2385    AATTTACCCAACGTTCCTTGGTAGTTGACCATGTGCGGCTGCTCCATTTCATGACGCCCG 1680
ISU3927   .G.....C.....A..G..C.....A..T. 1680
ISU55     .....C.....C.....A..T. 1680
ISU22     .G.....C.....G..C.....T.....A..T. 1680
VR2332    .G.....C...C....G..C.....T.....A..T. 1680
ISU1894   .G.....C...C....G..C.....T.....A..T. 1680
ISU79     .G.....C.....A.G..C.....A..T. 1680
                                                ***<ORF4
VR2385    AGACCATGAGGTGGGCAACTGTTTTAGCCTGTCTTTTTACCATTCTGTTGGCAATTTGAA 1740
ISU3927   .A..T.....CG..... 1740
ISU55     .....G.....C..... 1740
ISU22     .....G..... 1740
VR2332    .....G..... 1740
ISU1894   .A.....G.....C..... 1740
ISU79     .....C.....G.T..... 1740
                +1>ORF5
VR2385    TGTTTAAGTATGTTGGGGAAATGCTTGACCGCGGGCTGTTGCTCGCAATTGCTTTTTTTA 1800
ISU3927   .....G..C.....G 1800
ISU55     .....A.....TC.....G 1800
ISU22     .....C.....G.....C...G 1800
VR2332    .....A.....G.....C...G 1800
ISU1894   .....G.....C...G 1800
ISU79     .....T.....G.....C...G 1800
```

FIG. 1E

```
VR2385   TGGTGTATCGTGCCGTCTTGTTTTGTTGCGCTCGTCAGCGCCAACGGGAACAGCGGCTCA 1860
ISU3927  ..............TC...C...C.................AAC.G....---..C 1860
ISU55    ..................G.............C.........A.C.G..A.A....T 1860
ISU22    ....T...........TC.......C..T.....C........G.AAC.G....A....C 1860
VR2332   ..............TC.......C..T.....C..A.....G.AACG.....A....C 1860
ISU1894  ..............TC.......C..T.....C..A.....G..CC......A....C 1860
ISU79    ..............TC......AC..T.....C.GA....C..A.C......A....T 1860

VR2385   AATTTACAGCTGATTTACAACTTGACGCTATGTGAGCTGAATGGCACAGATTGGCTAGCT 1920
ISU3927  C........T.......T...C.............................C.....G... 1917
ISU55    C........T.......T.....................................T..... 1920
ISU22    C..C.................................................... 1920
VR2332   C..C.................................................... 1920
ISU1894  C..C.................................................... 1920
ISU79    C..C.G..AT.............................................. 1920

VR2385   AATAAATTTGACTGGGCAGTGGAGTGTTTTGTCATTTTTCCTGTGTTGACTCACATTGTC 1980
ISU3927  ...........T.......A..............C.....C.................T 1977
ISU55    GG.G.................................................... 1980
ISU22    ...........T.......A..............C.....C..T.............. 1980
VR2332   ..C........T.......A..............C.....C..T.............. 1980
ISU1894  G..........T.......A..............C.....C..T.............. 1980
ISU79    G...G......T.......A.C............C........T.............. 1980

VR2385   TCTTATGGTGCCCTCACTACTAGCCATTTCCTTGACACAGTCGGTCTGGTCACTGTGTCT 2040
ISU3927  ..C........A.....C..C.............................T......... 2037
ISU55    ..C.............C..C.................................... 2040
ISU22    ..C.................C.....................C.T.A.......... 2040
VR2332   ..C.................C.....................C.T.A.......... 2040
ISU1894  ..C.......................C..............CCT.A.......... 2040
ISU79    ..C.....C........C..C.................A.T.C.T.A.......... 2040

VR2385   ACCGCTGGGTTTGTTCACGGGCGGTATGTTCTGAGTAGCATGTACGCGGTCTGTGCCCTG 2100
ISU3927  .....C......CA.............................C.............. 2097
ISU55    .....C..C...TCC.............................C.............. 2100
ISU22    .....C.......................C..........C.............. 2100
VR2332   .....C.......................C..A........C.............. 2100
ISU1894  .....C.......................C..A........C.............. 2100
ISU79    .....C.......................C..A........C.............. 2100

VR2385   GCTGCGTTGATTTGCTTCGTCATTAGGCTTGCGAAGAATTGCATGTCCTGGCGCTACTCA 2160
ISU3927  ........T..................T..........C.................T 2157
ISU55    ............................T..A.......................... 2160
ISU22    ..........C.......A....C...T....A.....................G 2160
VR2332   ..........C..............T....A.....................G.G 2160
ISU1894  ........AGC..............T....A.....................T..G 2160
ISU79    ..........C..............T...T.......................... 2160
```

FIG. 1F

```
VR2385    TGTACCAGATATACCAACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTCG  2220
ISU3927   .....................C.....................................A  2217
ISU55     ..............................................................  2220
ISU22     .....................C.....................................  2220
VR2332    ..............................................................  2220
ISU1894   ..................................................AT...........  2220
ISU79     .....T.....................T.....................................  2220

VR2385    CCTGTCATCATAGAGAAAAGGGGCAAAGTTGAGGTCGAAGGTCACCTGATCGACCTCAAA  2280
ISU3927   ..................G....T..G.................T..........A...  2277
ISU55     .....................T.....................T.............G  2280
ISU22     ..........................................T...............  2280
VR2332    ..........................................T...............  2280
ISU1894   ..........................................T...............  2280
ISU79     ..................G.........................T.........T.....  2280

VR2385    AGAGTTGTGCTTGATGGTTCCGCGGCTACCCCTGTAACCAGAGTTTCAGCGGAACAATGG  2340
ISU3927   .A.........................A......T.........................  2337
ISU55     .........................A......A......A...........G......  2340
VR2385    ..............................................................  2340
ISU22     .....................T...A......A.....................  2340
VR2332    .....................T...A......A.....................  2340
ISU1894   ..............C.......T...A......A.....................  2340
ISU79     .....................T...A......A.....................  2340
                    ***<ORF5
VR2385    AGTCGTCCTTAG  2352
ISU3927   G.......C...  2349
ISU55     G...........  2352
VR2385    ............  2352
ISU22     G...........  2352
VR2332    G...........  2352
ISU1894   G...........  2352
ISU79     G...........  2352
```

FIG. 1G

A

```
VR2385  MKWGLC--K----AFLTKLAN-FLWMLSRSSWCPLLISLYFWPFCLASPSQVGWWSFASDWFAPRYSVRALPFTLSNYRRSYEAFLSQCQ  83
ISU22   ....P.--.-----.....................P.....................................................  83
ISU79   ....P.--.-----.....................P.....................................................  83
ISU55   .....--.-----.....S.....................P.....................................................  83
ISU1894 ....P.--.-----.....................S.....P.....................................................  83
ISU3927 .Q..P.--.-----....RSV.....................S.....LPA.....................................................  83
VR2332  ....P.--.-----.....................P.....................................................  83
LV      .Q..H.GV.SASCSWTPS.SSLLV.LI------.PF..---.Y..G....D.Y...F.E....F.....P.....GL.PN.R  80

VR2385  VDIPTWGTKHPLGMLWHHKVSTLIDEMVSRRMYRIMEKAGQAAWKQVVSEATLSRISSLDVVAHFQHLAAIEAETCKYLASRLPMLHMLR  173
ISU22   .....N..  173
ISU79   .....N..  173
ISU55   .....N..  173
ISU1894 .....N..  173
ISU3927 ....A...R.....G.....N..  173
VR2332  .....N..  173
LV      P.V.QFAV.....F..MR..H.....I.QT..HS.....G....TKL.G..I.T.....V..DS.RF.S...V..KN.A  170

VR2385  MTGSNVTIVYNSTLNQVFAVFPTPGSRPKLHDFQQWLIAVHSSIFSSVAASCTLFVVLWLRVPMLRTVFGFRWLGAIFLSNSR  256
ISU22   .....I.....I.S.....S..  256
ISU79   .....R..I.....I.....-  255
ISU55   .....L.I.....IA.....Q  256
ISU1894 I.....SE...I.....G..P...W  256
ISU3927 I.....H.....I.....S.....M...S.....P.S.W  256
VR2332  .....I.....Q  256
LV      V--G..SLQ..T..DR.ELI.....T....T..R....S..A.....S.V...I.....I.A..Y....H.PT.THH.S--  249
```

FIG. 2A

SUBSTITUTE SHEET (RULE 26)

B

```
VR2385   MANSCTFLYIFLCCSFLYSFCCAVVAGSNATYCFWFPLVRGNFSFELTVNYTVCPPCLTRQAAAEAYEPGRSLWCRIGHDRCGEDDHDEL  90
ISU55    ..........................................................T.............Y.....  90
ISU79    .....A..H.........L.........T.................I.............................Y.....  90
ISU1894  .V......H..............A.....................M.................V.....Y.....  90
ISU22    .V......H.....................T.F.....................................I.....Y.....  90
ISU3927  ........H.L...........V..TDA...F..............M.................QI...N.......N.....  90
VR2332   .V......H.....................T.....................................T.I.....Y.....  90
LV       ..HQ.ARFHF...GFIC.LVHS.LASN.SS.L......AH..T.....I...I.M..S.S...RQRL....NM..K......E.R.....  90

VR2385   GFVVPSGLSSEGHLTSAYAWLASLSFSYTTQFHPEIFGIGNVSRVYVDIKHQFICAVHDGQNTTLPHHDNISAVLQTYYQHQVDGGNWFH  180
ISU55    ................V.....F......A.............Q.....R.............A...R.......F.....  180
ISU79    ..MI............V.....F......A.....................L...E.........R.......F.....  180
ISU1894  ..MI.P..........V.....F......A.....................L...E.....A...R.......F.....  180
ISU22    ..M..P.F........V.....F......A.....................L...E.........R.......F.....  180
ISU3927  ..T..P...K.V....V.....F......A.............K.....N..L.............R.......F.....  180
VR2332   ..MI.P.........GV.....F......A.....................L...E.........R.......F.....  180
LV       LMSI...Y-DNLK.EGY.....F.....AA.....L.........F..KR......E...H.S.VSTGH....LYAA..H..I.....  179

VR2385   LEWVRPFFSSWLVLNVSWFLRRSPASHVSVRVFQTSRPTPPQRQALLSSKTSV--ALGIATRPLRRFA--------KS--LSAARR  254
ISU55    ...L.............................L..L.....................--.....--------.---..VV..  254
ISU79    ...L.....................N......L..L.....................--.....--------.---...V..  254
ISU1894  ...L..............Y......N.A....L.IL...L.....................--.....--------.---...V..  254
ISU22    ...L.....................N......L.IL.....................--.....--------.---...V..  254
ISU3927  ...L.....................................RQ.IS...R..A---...M.....--------.--.....  254
VR2332   ...L.....................N......L.IL.....................--.....--------.---...V..  254
LV       ...L..L........I........V.P..R.IY.IL...R.RLPVSW.FR..IVSD.TGSQQRK.K.PSESRPNVV.PSV.PSTS.  265
```

FIG. 2B

C

```
VR2385   MGASLLFLLVGFKCLLVSQAFACKPCFSSSLSDIKTNTTAAAGFAVLQDISCLRHRN-SASE---AIRKVPQCRTAIGTPVYITVTANVT  86
VR2332   .AS.....V.....................A.........S.............D-....---....I.............V.I.....  86
ISU55    .A.............................A.........................Y..-....---.F..I..........M.........  86
ISU1894  .AS.....M.....................A.........S.............-....---....I.............I.....  86
ISU22    .AS.....M.....................A.........S.......G.....D-....---....I.............I.....  86
ISU79    .A......M.....................A.........S.............-....---....I..........M...I.....  86
ISU3927  .A..........E.....................N.......G....G.-.TT.---.F..I.............I.....  86
LV       .A.AT..F.A.AQHIM..E.........TH....E.........M.....N.F.PHGV..AQEKISFG.SS...E.V...Q...I.....  90

VR2385   DENYLHSSDLLMLSSCLFYASEMSEKGFKVVFGNVSGIVAVCVNFTSYVQHVKEFTQRSLVVDH-VRLLHFMTPETMRWATVLACLFTIL  175
VR2332   ..........................................................-.....................A..  175
ISU55    ..........................................................-.....................A..  175
ISU1894  ..................................................R.......-.....................A..  175
ISU22    ..................................................R.......-.....................A..  175
ISU79    ..........................E.......................R......M..-.....................A..  175
ISU3927  ....................................T....I.....................-.....................A..  175
LV       ..S..YNA......A..............I......V.SA....D..A..TQH..QHHL.IDHI.....L..SA.....TI....A..  180

VR2385   LAI  178
VR2332   ...  178
ISU55    ...  178
ISU1894  ...  178
ISU22    ...  178
ISU79    ...  178
ISU3927  ...  178
LV       ...  183
```

FIG. 2C

SUBSTITUTE SHEET (RULE 26)

D

```
                                                  *******
VR2385   MLGKCLTAGCCSQLLFLWCIVPSCFVA--LVSANGNSGSNLQLIYNLTLCELNGTDWLANKFDWAVECFVIFPVLTHIVSYGALTTSHFL  88
VR2332   ..E.........R..S......F..AV--.AN.SND.S.H.....................................S.....................  88
ISU55    .........Y..S..........W...--.A...SSNS.H.....................GE.....................................  88
ISU1894  ............R..S......F..AV--.AN.SA..S.H.....................D.....  .S.....................L.  88
ISU79    ............R..S..F...F..AV--.A..SNS.S.H.....................................S.....................  88
ISU22    .......V....R..S......F..TV--.AD.HS..S.H.....................DR......S.....................  88
ISU3927  ............RS........F.LA.--.....-..S.H.....................................S.....................  87
LV       .RCSHKLGRFLTPHSCF.WLFLL.TGLSWSF.D..GDS.TY.Y.....I.........SSH.G....T..LY..A...L.L.F.....F  90

VR2385   DTVGLVTVSTAGFVHGRYVLSSMYAVCALAALICFVIRLAKNCMSWRYSCTRYTNFLLDTKGRLYRWRSPVIIEKRGKVEVEGHLIDLKR  178
VR2332   ...A.....................I.........T.....F.........A.....................................  178
ISU55    .............S........I................FT.....................................  178
ISU1894  ...A.....................I.........A.....F.....................................H.....  178
ISU79    ..IA.....................I.........T.....FV.....................................  178
ISU22    ...A.....................I.........T..I..F.....................................  178
ISU3927  .............H........I........F......F.....................................G.............K  177
LV       .AL..GA.......G.....C.V.GA..F..FV.....A.....AC..AR..F...IV.DR..VH..K..IVV..L..A..D.N.VTI.H  180

VR2385   VVLDGSAATPVTRVSAEQWSRP  200
VR2332   ......V...I........G..  200
ISU55    ..........I.K......G..  200
ISU1894  ......V...I........G..  200
ISU79    ......V...I........G..  200
ISU22    ......V...I........G..  200
ISU3927  ..........L........G..  199
LV       ...E.VK.Q.L..T.....EA-  201
```

FIG. 2D

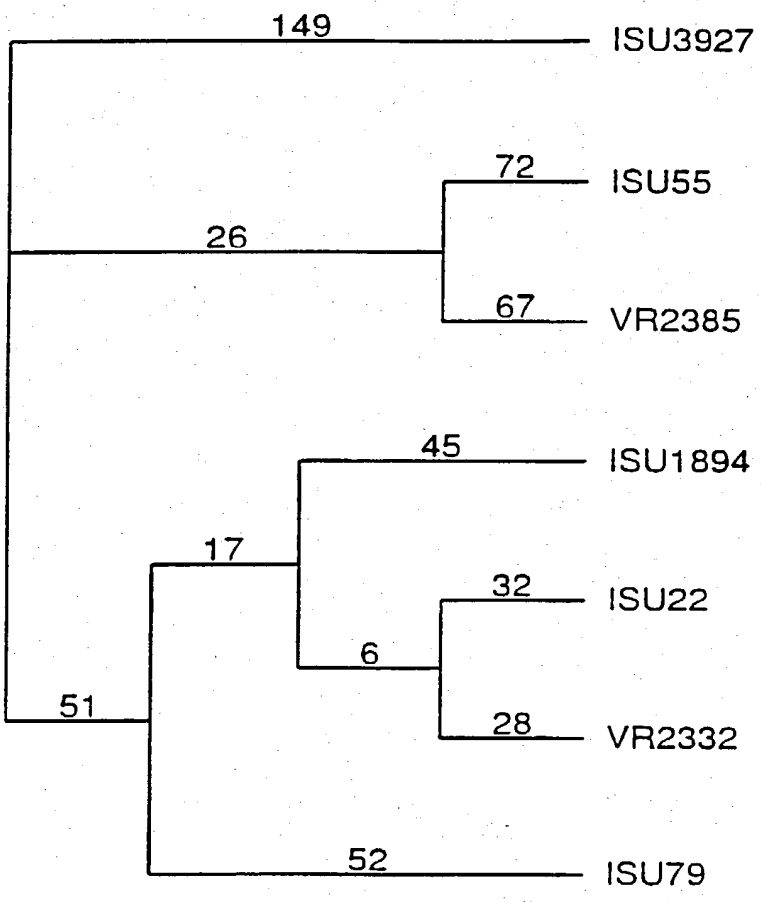
149
ISU3927
72
ISU55
26
67
VR2385
45
ISU1894
17
32
ISU22
6
51
28
VR2332
52
ISU79

FIG. 3

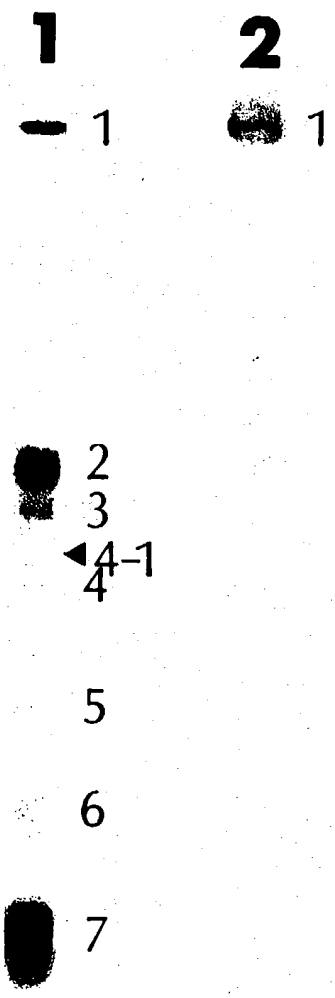

FIG.4

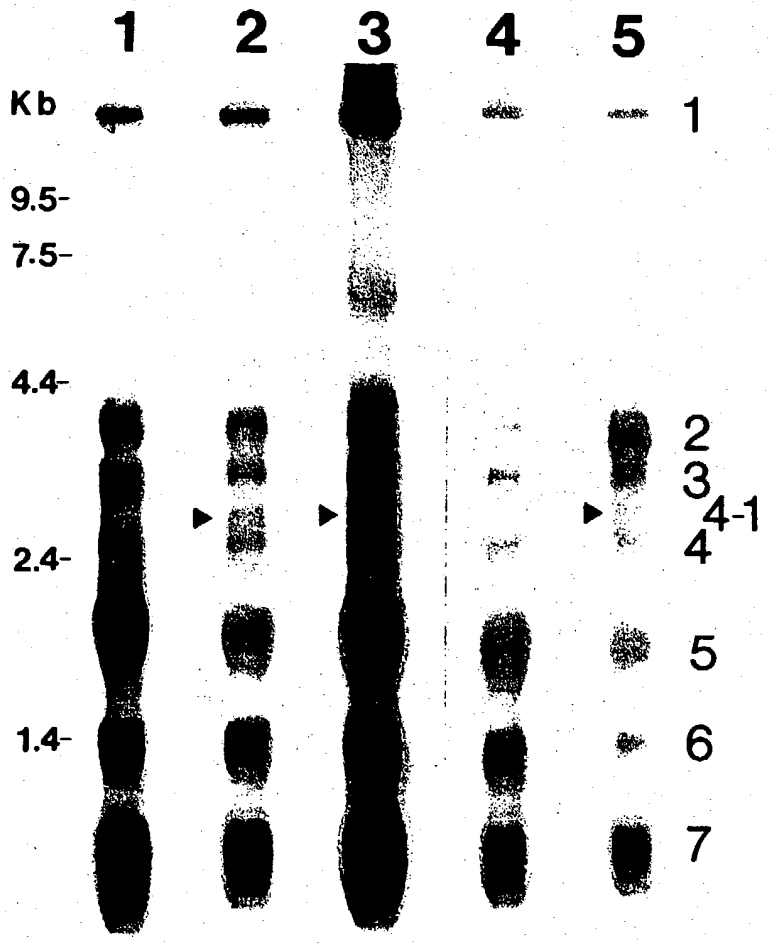

FIG.5

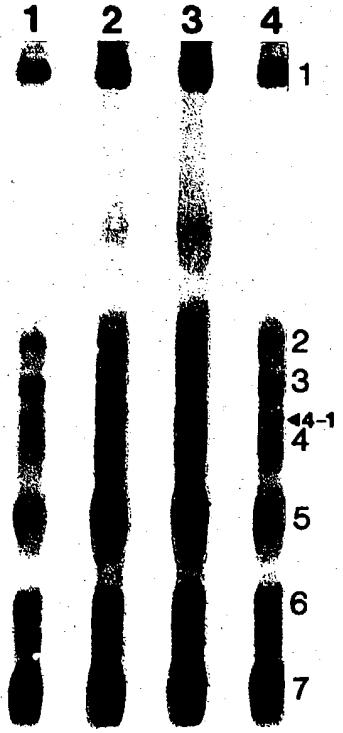

FIG.6A

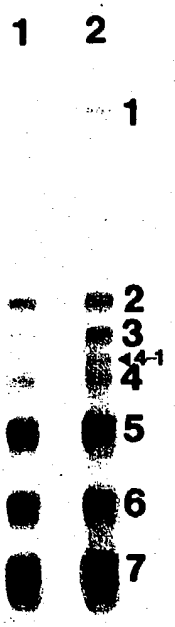

FIG.6B

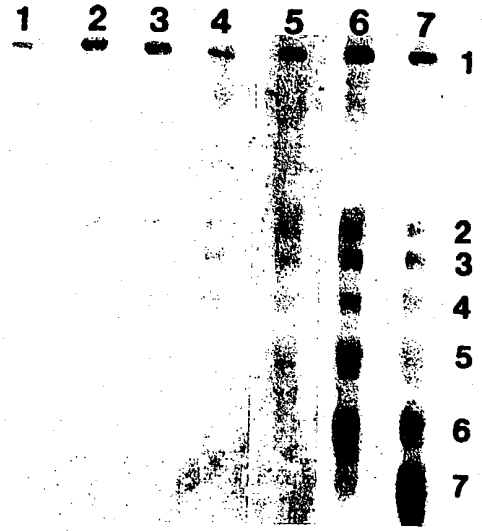

FIG.7A

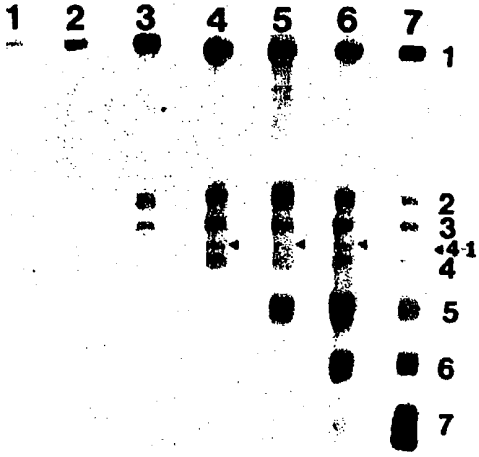

FIG.7B

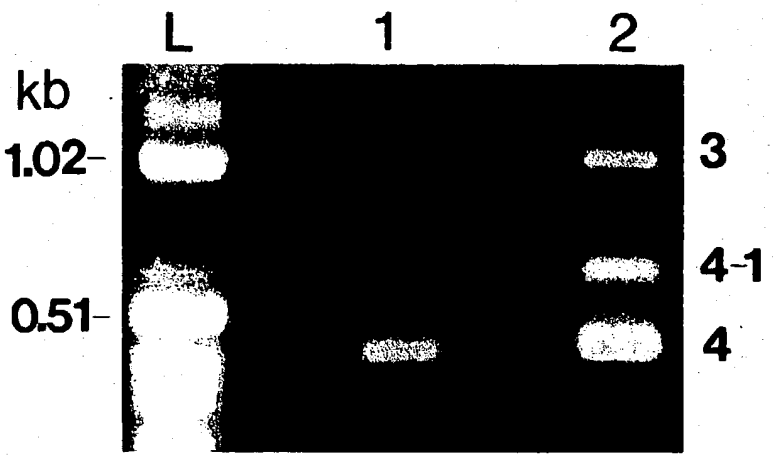

FIG.8A

	-89	+1>ORF3
ISU-1894-mRNA3	GUAACC	...AUG
ISU-79-mRNA3	GUAACC	...AUG
	-236	+1>ORF4
ISU-1894	UUGACu	...AUG
ISU-79-mRNA4-1	UUGACC	...AUG
	-10	+1>ORF4
ISU-1894-mRNA4	UUCACC	...AUG
ISU-79-mRNA4	UUCACC	...AUG

FIG.8B

```
                                    -26 (mRNA2)                    +1>ORF2
ISU79   GTTTTATTTCCCTCCGGGCCCTGTCATTGAACCAACTTTAGGCCTGAATTGAAATGAAAT  60
ISU1894 .....................C.....______.............................

ISU79   GGGGTCCATGCAAAGCCTTTTTGACAAAATTGGCCAACTTTTTGTGGATGCTTTCACGGA 120
ISU1894 .......G.....................................................  120

ISU79   GTTCTTGGTGTCCATTGTTGATATCATTATATTCTTGGCCATTTTGTTTGGCTTCACCAT 180
ISU1894 ....C.....................C.....T.............................  180

ISU79   CGCCGGTTGGCTGGTGGTCTTTTGCATCAGATTGGTTTGCTCCGCGATACTCCGTACGCG 240
ISU1894 .....................................................................  240

ISU79   CCCTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTTCCCAGTGCC 300
ISU1894 ...................................................T.A.....  300

ISU79   AAGTGGACATTCCCACCTGGGGAACTAAACATCCTTTGGGGATGTTTTGGCACCATAAGG 360
ISU1894 .....................................................................  360

ISU79   TGTCAACCCTGATTGATGAGATGGTGTCGCGTCGAATGTACCGCATCATGGAAAAAGCAG 420
ISU1894 .....................A.....................C.....................  420

ISU79   GACAGGCTGCCTGGAAACAGGTGGTGAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGG 480
ISU1894 .G...................................................C.....  480

ISU79   ATGTGGTGGCTCATTTTCAGCATCTTGCCGCCATCGAAGCCGAGACCTGTAAATATTTGG 540
ISU1894 .....................................T.....T.....  540
                                                          -89(mRNA3)
ISU79   CCTCCCGGCTGCCCATGCTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGT 600
ISU1894 .....................T.....A.....______.....  600

ISU79   ATAATAGTACTTTGAATCGGGTGTTTGCTATTTTCCCAACCCCTGGTTCCCGGCCAAAGC 660
ISU1894 .....C.G.G.A.....  660
                    +1>ORF3
ISU79   TTCATGACTTTCAGCAATGGCTAATAGCTGTGCATTCCTCCATATTTTCCTCTGTTGCAG 720
ISU1894 .....T.....•••.T.....A.....  720

ISU79   CTTCTTGTACTCTCTTTGTTGTGCTGTGGTTGCGGGTTCCAATACTACGTACTGTTTTTG 780
ISU1894 .....T.....C.....G.....  780
                                                  ***<ORF2(ISU79)
ISU79   GTTTCCGCTGGTTAGGGGCAATTTTTCTTTCGAACTCATAGTGAATTACACGGTGTGCCC 840
ISU1894 .....G.....C.....G.....T..  840
                                                  ***<ORF2(ISU1894)
```

FIG. 9A

```
ISU79   ACCTTGCCTCACCCGGCAAGCAGCCGCAGAGGCCTACGAACCCGGTAGGTCTCTTTGGTG 900
ISU1894 ..............................T.......................... 900

ISU79   CAGGATAGGGTACGATCGATGTGGAGAGGACGACCATGACGAGCTAGGGTTTATGATACC 960
IAU1894 ............T..C........G.........T.......................... 960
                                    -236(ISU79 mRNA4-1)
ISU79   GTCTGGCCTCTCCAGCGAAGGCCACTTGACCAGTGTTTACGCCTGGTTGGCGTTCTTGTC1020
ISU1894 .C............................T.....................T.....1020

ISU79   CTTCAGCTACACGGCCCAGTTCCACCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGT1080
ISU1894 ........................T.....................................1080
          +1>ORF4-1
ISU79   TTATGTTGACATCAAACATCAACTCATCTGCGCCGAACATGACGGGCAGAACACCACCTT1140
ISU1894 ............................T................A......G.......1140

ISU79   GCCTCGTCATGACAACATTTCGGCCGTGTTTCAGACCTATTACCAACATCAAGTCGACGG1200
ISU1894 ..............T.....A.....................................T..1200
                    -10(mRNA4)+1>ORF4
ISU79   CGGCAATTGGTTTCACCTAGAATGGCTGCGTCCCTTCTTTTCCTCATGGTTGGTTTTAAA1260
ISU1894 ..........._____.........T.............................1260
                          ***<ORF4-1
ISU79   TGTCTCTTGGTTTCTCAGGCGTTCGCCTGCAAACCATGTTTCAGTTCGAGTCTTGCAGAC1320
ISU1894 ...........A.........................C.....................T1320

ISU79   ATTAAGACCAACACCACCGCAGCGGCAAGCTTTGCTGTCCTCCAAGACATCAGTTGCCTT1380
ISU1894 ..............T.....................................................1380
                                                                    **
ISU79   AGGCATCGCAACTCGGCCTCTGAGGCGATTCGCAAAATCCCTCAGTGCCGTACGGCGATA1440
ISU1894 .....................................................................1440
        *<ORF3
ISU79   GGGACACCTATGTATATTACCATCACAGCCAATGTGACAGATGAAAATTATTTACATTCT1500
ISU1894 ........CG.................................G..............1500

ISU79   TCTGATCTCCTCATGCTCTCTTCTTGCCTTTTCTATGCTTCTGAGATGAGTGAAAAGGGA1560
ISU1894 ................T..............C...................A...1560

ISU79   TTTGAGGTGGTTTTTGGCAATGTGTCAGGCATCGTGGCTGTGTGTGTCAATTTTACCAGC1620
ISU1894 ..CA.......A.....................................................1620
                                                        -105
ISU79   TACGTTCAACATGTCAGGGAGTTTACCCAACGCTCCTTGATGGTCGACCATGTGCGGCTG1680
ISU1894 .....C.............................C..G...._____.........T..1680
                          -70              -55(ISU79 mRNA5)
ISU79   CTCCATTTCATGACACCTGAGACCATGAGGTGGGCAACCGTTTTAGCCTGTCTTTTTGCT1740
ISU1894 ..................__A__.............T...................C1740
                          -70(ISU1894 mRNA5)
```

FIG. 9B

```
                    ***<ORF4          +1>ORF5
ISU79    ATTCTGTTGGCAATTTGAATGTTTAAGTATGTTGGGGAAATGCTTGACCGTGGGCTGTTG1800
ISU1894  ......C.....................................C.........1800

ISU79    CTCGCGATTGCTTTCTTTGTGGTGTATCGTGCCGTTCTGTTTTACTGTGCTCGCCGACGC1860
ISU1894  .....................................G...........A....1860

ISU79    CCACAGCAACAGCAGCTCTCATCTGCAATTGATTTACAACTTGACGCTATGTGAGCTGAA1920
ISU1894  .AG.GC...........C.....A..GC.............................1920

ISU79    TGGCACAGATTGGCTAGCTGATAGATTTGATTGGGCAGTGGAGAGCTTTGTCATCTTTCC1980
ISU1894  ......................A.....................T.............1980

ISU79    TGTTTTGACTCACATTGTCTCCTATGGCGCCCTCACCACCAGCCATTTCCTTGACACAAT2040
ISU1894  C.........................T........T..T......C...........G.2040

ISU79    TGCTTTAGTCACTGTGTCTACCGCCGGGTTTGTTCACGGGCGGTATGTCCTAAGTAGCAT2100
ISU1894  C..C.....................................................2100

ISU79    CTACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTCATTAGGTTTGTGAAGAATTG2160
ISU1894  ........................AG.....................CA.........2160

ISU79    CATGTCCTGGCGCTACTCATGTACTAGATATACCAACTTTCTTCTGGATACTAAGGGCAG2220
ISU1894  ..............T..G.....C.......................C...........2220

ISU79    ACTCTATCGTTGGCGGTCGCCTGTCATCATAGAGAAGAGGGGCAAAGTTGAGGTCGAAGG2280
ISU1894  ..............AT.....................A.......................2280
                                                    -32       -23(mRNA6)
ISU79    TCATCTGATCGATCTCAAAAGAGTTGTGCTTGATGGTTCCGTGGCAACCCCTATAACCAG2340
ISU1894  ............C.....................C.........______....._____..2340
                        +1>ORF6       ***<ORF5
ISU79    AGTTTCAGCGGAACAATGGGGTCGTCCTTAGATGACTTCTGTTATGATAGTACGGCTCCA2400
ISU1894  .....................................CC..................2400

ISU79    CAAAAGGTGCTTTTGGCATTTTCTATTACCTACACGCCAGTAATGATATATGCCCTAAAG2460
ISU1894  ................G.....................G.................2460

ISU79    GTGAGTCGCGGCCGACTGCTAGGGCTTCTGCACCTTTTGATTTTCCTGAACTGTGCTTTC2520
ISU1894  .....................................C........T.........2520

ISU79    ACCTTCGGGTACATGACATTCATGCACTTTCAGAGTACAAATAAGGTCGCGCTCACTATG2580
ISU1894  ..................G.....................................2580
```

FIG. 9C

```
ISU79   GGAGCAGTAGTTGCACTCCTTTGGGGGGTGTACTCAGCCATAGAAACCTGGAAATTCATC2640
ISU1894 ..................................................................2640

ISU79   ACCTCCAGATGCCGTTTGTGCTTGCTAGGCCGCAAGTACATTCTGGCCCCTGCCCACCAC2700
ISU1894 ..................................................................2700
                                                        -129
ISU79   GTTGAAAGTGCCGCAGGCTTTCATCCGATTGCGGCAAATGATAACCACGCATTTGTCGTC2760
ISU1894 ...........................................      ..........2760

ISU79   CGGCGTCCCGGCTCCACTACGGTCAACGGCACATTGGTGCCCGGGTTGAAAAGCCTCGTG2820
ISU1894 ..................................................................2820
                                                -15 (mRNA7)    +1>ORF7***<ORF6
ISU79   TTGGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCTTGTCAAATATGCCAAATAA2880
ISU1894 ...................................      ..........................2880

ISU79   CAACGGCAAGCAGCAGAAGAGAAAGAAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCA2940
ISU1894 ..................................................................2940

ISU79   GATGCTGGGTAAGATCATCGCCCAGCAAAACCAGTCTAGAGGCAAGGGACCGGGAAAGAA3000
ISU1894 ..................T.............C.....................3000

ISU79   AAATAAGAAGAAAAACCCGGAGAAGCCCCATTTTCCTCTAGCGACTGAAGATGATGTCAG3060
ISU1894 ...C..............................................................3060

ISU79   ACATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTCGTCAATCCAAACTGCCTTTAA3120
ISU1894 .........C.....................................G..C........3120

ISU79   TCAAGGCGCTGGGACTTGCACCCTGTCAGATTCAGGGAGGATAAGTTACACTGTGGAGTT3180
ISU1894 ..................................................................3180
                                                                      **
ISU79   TAGTTTGCCTACGCATCATACTGTGCGCTTGATCCGCGTCACAGCATCACCCTCAGCATG3240
ISU1894 .........A........................................................3240
        *<ORF7
ISU79   ATGGGCTGGCATTCTTGAGGCATCCCAGTGTTTGAATTGGAAGAATGCGTGGT 3293
ISU1894 ..................................................... 3293
```

FIG. 9D

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/08962

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.72, 23.1; 530/350; 424/186.1, 159.1; 514/44; 530/389.4, 388.3; 435/5, 69.3, 91.53, 91.2, 91.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, STN, CAPLUS, CAB, MEDLINE, BIOSIS, WPIDS
search terms: Mystery Swine, Blue Ear, PRRSV, SIRS, PEARS, swine infertility, Lelystad

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0,595,436 A2 (SOLWAY ANIMAL HEALTH, INC.) 04 MAY 1994, see entire document, especially ages 6-14 and pages 24-28.	1, 4-12, 13-21
X ----- Y	MOROZOV et al. Sequence analysis of open reading frames (ORFs) 2 to 4 of a U.S. isolate of porcine reproductive and respiratory syndrome virus. Arch. Virology. 1995, Vol. 140, pages 1313-1319, see entire document.	1, 5-9 ----- 1-21
Y	NELSON et al. Differential of U.S. and European Isolate of Porcine Reproductive and Respiratory Syndrome Virus by Monoclonal Antibodies. Journal of Clinical Microbiology. December. 1993, Vol. 31, No. 12, pages 3184-3189, see entire document, especially pages 3187-3188.	1-21

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 03 SEPTEMBER 1996	Date of mailing of the international search report 04 OCT 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DANNA LEE Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/08962

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GB 2,282,811 A (CYANAMID IBRECA S A) 19 April 1995, see entire document, especially pages 15-19.	1-21
Y	WO 93/07898 A1 (AKZO N.V.) 29 April 1993, see entire document, especially pages 5-11.	1-21
Y	WO 92/21375 A1 (STICHTING CENTRAAL DIERGENEESKUNDIG INSTITUUT) 10 December 1992, see entire document, pages 2-8, 29, 40-43.	1-21
A	GOYAL, S.M. Porcine reproductive and respiratory syndrome. J. Vet. Diagn. Invest. 1993, Vol. 5, pages 656-664, see entire document.	1-21
A	DOMINGO et al. New observations on antigenic diversification of RNA viruses. Antigenic variation is not dependent on immune selection. Journal of General Virology. 1993, Vol. 74, pages 2039-2045, see entire document.	1-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/08962

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12N 15/40, 15/11, C07K 14/08; C12P 21/02, 19/34; C12Q 1/70; G01N 33/569; A61K 39/12, 39/42, 48/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.72, 23.1; 530/350; 424/186.1, 159.1; 514/44; 530/389.4, 388.3; 435/5, 69.3, 91.53, 91.2, 91.1